The organisational impact of key performance indicators : their effect on English and Scottish schools

Wiggins, Andrew

The organisational impact of Key Performance Indicators: Their effect on English and Scottish schools

Andrew Wiggins
PhD

University of Durham
School of Education

2002

Abstract

The use of key performance indicators has dramatically increased over the last twenty or so years. Their introduction has been largely justified on the grounds that high stakes proxy indicators (test and exam results), increase accountability which will in turn 'drive up' overall performance. Whilst there has been some evidence that measured standards have risen, there have been associated costs, as well as claims of unintended or dysfunctional consequences.

The place of KPIs within the broader organisational and managerial context was firstly considered, with particular reference to generic management and accounting theory. Secondly, the results of a questionnaire survey of 162 heads and teachers in England and Scotland was reported. The key findings included evidence that KPIs had a narrowing effect on the curriculum, that they tended to undermine heads and teachers, and encouraged a blame culture. There was also evidence that they caused schools to concentrate on targets at the expense of other important objectives, as well as concentrating on 'boarder-line' pupils.

English primary schools reported far more dysfunctional behaviour due to KPIs, than did their Scottish counter parts. This was attributed to league tables which Scottish primaries do not have. At the secondary level the results were similar, tables are published for secondary schools in both countries. There was widespread support for changes to the KPIs, including the use of a wider range of measures and 'value added' indicators, as well as discontinuing league tables.

Improvements to the KPI systems were discussed, including the use of 'balanced score card' systems; however, it was argued that such technical changes need to be accompanied by more fundamental organisational changes. There needs to be 'top down' leadership, which devolves trust and responsibility, rather than blame and accountability; and one that will develop and nurture a true learning culture throughout the education system.

Acknowledgments

For: Chris, Charlotte and Peter

Without whose support and understanding this would not have been possible

There were many notable contributors. Firstly, Peter Tymms, whose insightful and patient guidance and inspiration, both identified and drew out many of the important issues considered by this thesis. And secondly, the many people from both England and Scotland, who generously gave time to offer their thoughts and experiences.

Table of Contents

List of Tables

Table of Figures

Table of survey questions

Chapter 1

Introduction

Not everything that counts can be counted, and not everything that can be counted counts

(Einstein)

Rationale

The last twenty or so years has seen a dramatic rise in the use of performance indicators (PIs), both in education and in the public sector more generally. Their use has been largely justified on the grounds that they will increase accountability, which will in turn improve management and performance; however, little consideration has been given to their broader organisational impact. Ouston *et al* (1998) point out that at the policy level accountability based on PIs is viewed as unproblematic, but that in practice the 'story' is far more complex:

> In recent years, educational accountability has become a widely discussed...It is often presented as quite unproblematic: Its meaning and its implementation are seen as entirely straightforward. However, on examination such confidence in the concept cannot be supported.
> (p. 111)

This thesis looks at the 'story' behind performance indicators, their effect on the organisation and management of schools, and in particular how they may encourage dysfunctional behaviour. The main indicators considered here are those which have come to be referred to as Key Performance Indicators[1] (KPIs). In England this includes, Key Stage 2 test results and GCSEs, and in Scotland National 5-14 testing and Standard grades.

Although quite a lot has been written about the effects, much of this has been based on individual judgements and observations, or as an adjunct to other research. Very little specific detailed research on the main issues has been carried out. This thesis aims to contribute to this area in two main ways; firstly by drawing on theories and evidence from other fields, and secondly by carrying out a specific research programme.

The main elements of the thesis

There are four main elements to the thesis. These aim to include a wide range of theory and experiences from other sectors, as well as reporting on the detailed research programme.

A broad organisational perspective: The first element looks at the place of PIs in the broader organisational and management context. The reason for this approach is the argument that PIs do not exist in isolation, but are a part of, and interact with other aspects of the management and organisation. From this perspective, and indeed for much of the thesis, substantial reference is made to 'generic management theory'. However, there has to date been relatively little use of such

---

[1] Originally the term Public Performance Indicators (PPIs) was used by this research, however this was changed to KPIs as this became widely accepted. All references to PPIs have been changed to KPIs in this thesis.

theory in educational management research, although examples do exist of its effective use. For instance, Cowie (2000) uses Mintzberg's organisational configurations[2] to help develop a theoretical model to explore the relationships between heads and their Education Authority (EA).

Business accounting: The business accounting world has been grappling for hundreds or even thousands of years with the problem of measuring performance. Like education, this performance is ultimately very much dependent on the commitment and creativity of individuals, as well as the effectiveness of the organisation and its management. Therefore, it is argued that much can be learnt from the experiences of the broader business accounting sector, as well as the more specific behavioural field. A number of theorists, for example, Broadbent *et al* (1996) have applied aspects of behavioural accounting theory (eg. Principle – Agent theory) to education, and these do provide useful models to help explore some of the effects of PI systems on schools.

The wider public sector: The substantial changes in education have been mirrored in many (if not all) other areas of the public sector. As with education these sectors have similar difficulties in measuring what is important, and their ultimate performance (ie. not just what is measured) is very much dependent on the commitment and dedication of the people providing the service. Therefore, it would seem that much can be learnt from these other sectors, this is particularly so, when considering that ultimately the 'whole' public sector is managed by the same body (ie. the government). A number of theorists, for example Smith (1995a) have developed theoretical models (*unintended consequences of PIs*) which are universally applicable to much of the public sector. Furthermore, there are increasing calls for 'joined up' government and consequently opportunities for increased inter-sector learning. In addition, joint research initiatives, such as the Durham Evidence Based conferences provide good examples of learning from each other.

Research programme: In spite of the importance of the fundamental issue of how PIs affect organisational behaviour, there has to date been very little detailed and specific research aimed at schools. Details of research carried out with primary and secondary schools in England and Scotland during 1999 / 2000 are reported and discussed in the fourth element of the thesis.

Review of existing research

To help establish and illustrate some of the key issues, a number of references to the behavioural effects of PIs are discussed below. These do not however provide a clear and unambiguous view of the effects of PIs; many of these references will be discussed further in the following chapters.

---

[2] Mintzberg 1993

The government's position has remained largely unchanged since the early 1990's. 'Their' view is that the key to the necessary educational improvements is a programme of regular testing, and the widespread publication of the results (Cabinet Office 1991). Indeed, the current Prime minister, in spite of possible reservations, and when in opposition the objection to league tables, states that the government have a duty to publish such data (Barnard 1999). Whilst few would disagree with the principle of free access to relevant data, much depends on how this is done, and what use is made of the results.

John Dunford, general secretary of the Secondary Heads Association (Carvel and Smithers 1998), points out that the league tables, which are rankings of the raw KPI data, distort priorities and encourage schools to concentrate on small groups of borderline[3] children. Ofsted's previous chief inspector, Mike Tomlinson, suggested that league tables discourage inclusion (Thornton 2001), and 'widen the gap' by encouraging schools to concentrate on more able pupils; on the other hand there is evidence of schools neglecting the most able children (Rafferty 1999). This neatly illustrates the importance of contextualizing such views and findings; in effect these three views suggest that league tables encourage schools to concentrate on children at the top, middle and bottom!

In terms of the curriculum, Earl *et al* (2001) found that the subjects which are tested (Literacy and Numeracy) were displacing other subjects. For example, there is evidence that the testing and league tables have encouraged schools to reduce the amount of Physical Education and sports (Grace 2000b), and this may well contribute to a decline in the overall physical health of the nation. Similar claims have been made for the aesthetic and spiritual well-being of children, with the reduction in time for subjects such as art, music and religious education.

James (2000) argues that the continual testing regime throughout childrens' school lives, discourages 'lifelong learning'. In a similar way, Pollard and Triggs (2000) found that the testing had caused children to move from being learning orientated, to performance orientated with a 'low tolerance of ambiguity'. Both of these will affect the overall educational potential of the nation. Hartley-Brewer (2001) points to the growing evidence that the high stakes testing regime can have a negative impact on childrens' mental health, with an increase in psychiatric referrals. The Professional Association of Teachers (in conjunction with the childrens' society) aptly named report, *Tested to destruction* (2000), chronicles the testing regime which runs throughout children's school lives, and describes some of the associated problems, in particular the stress they cause throughout the education system. Related to this, the pressure appears to have encouraged an ever

---

[3] Those close to reaching national targets

increasing number of teachers and schools to cheat to help 'their' pupils do well in the key stage tests (Smithers 2002).

As Earl *et al* (2001) point out the potential negative effects are not confined to the children. They raised concerns that the National Literacy and Numeracy Strategies (NLNS) which are designed to raise KPI performance are having a negative effect on schools and teachers, by creating a culture of dependency rather than initiative and creativity.

Goldstein and Spiegelhalter (1996) discuss the limitations of league tables, and in particular the difficulties the public face in making proper interpretations. Related to this Goldstein and Myers (1996), whilst broadly supporting the publication of data, propose a code of ethics for their use and production. In Scotland, Croxford (1999) points to the misleading nature of league tables based on raw scores, and suggests that possible future league tables based on school level targets would be even more misleading, because of the spurious impression of fairness. In a similar way in England, Henry (2001) found that proposed improvements to the performance tables in the form of value added scores may be misleading and mistrusted by heads. Whilst targets, tests and tables might be of use to schools and the public, Hallgarten (2002) discusses the use made by politicians, and suggests that it is time politics were taken out of the tables.

In the US where high stakes testing is quite widespread, Klein *et al* (2000) suggest that testing in Texas may hinder the development of literacy and numeracy skills; furthermore, they found evidence that they further disadvantaged racial minorities. Haney (2000) also in Texas, found that the exclusion of pupils with disabilities had increased, whilst having decreased in the rest of the US. And more generally in the US, Miron (1999) found that the high stakes testing systems encouraged charter schools[4] to 'cream skim' the best and easiest to deal with pupils.

Structure of thesis

The first chapter of the main body of the thesis, chapter 2, looks at some of the technical issues surrounding PI. Issues of accountability and the use of PIs in the management and organisation of schools is considered. The quality of data and the concept of proxy and 'high stakes' indicators are explored, along with how indicators are used at different levels of the system. Finally, the issue of 'who manages what' in the overall education process is discussed.

---

[4] Similar to City academies

Theory from other sectors is used throughout the thesis, and in chapter 3 some of the arguments for and against are explored. It is also argued that measuring business and educational performance provides similar challenges, and that the theory from both fields is converging. Leading on from this, chapter 4 uses generic management theory to look at the broader organisational context, within which performance indicators exist. A range of theoretical schools are considered, including the work of FW Taylor, Elton Mayo and the Total Quality Management approach.

Chapter 5 looks at the development of public sector management, including 'new public management', and the citizen's charter. From this the development of educational management is considered, including the influential 1988 Education Reform Act. The introduction of compulsory national testing, in both England and Scotland, and the issue of performance and league tables is also discussed.

Some of the research on the behavioural aspects of performance indicators in business is considered in chapter 6. Three key theorists in particular are discussed; Chris Argyris, Anthony Hopwood and David Otley. The effect of high stakes league tables on businesses is also considered, along with the seminal work of Johnson and Kaplan which questioned a number of long established accounting principles. Chapter 7 looks at the use of high stakes PIs in other areas of the public sector. Firstly, the important work of Peter Smith on the unintended consequences of PI systems, and then more specifically the use of PIs for health, railways and the police is discussed.

Chapter 8 looks at the design of the research which was carried out, and discusses whether a quantitative or qualitative approach should be taken. A number of possible research instruments are considered, and from this the final design is outlined. To help establish some of the important issues, preliminary research was undertaken, and this is discussed in chapter 9. A number of key hypotheses are also established. Chapter 10 discusses the methodology for the main research: this includes the question design and formats, distribution and response rates.

The results are discussed in three parts. Firstly, chapter 11 provides a descriptive analysis of all of the responses, in the form of graphs and statistics, accompanied by a general discussion on the particular issues. Secondly, chapter 12 describes a comparative analysis of the responses, using factors such as country, phase and type of post. This is carried out with a number of the questions, and t tests and Effect Sizes are used to calculate the significance. Finally, chapter 13, firstly reports on a follow up telephone survey to assess possible changes in the responses over time, and then reports on some of the relationships or degrees of correlation between the results.

Chapter 14 the conclusion. provides a summary of the chapters which make up the thesis. Then five key themes; *targets*, *accountability*, *complexity*, *the right measures and broader organisational context*: which run throughout the thesis and cut across the theory and research are discussed. A brief review of how KPI systems may develop in the future is carried out. followed by a critique of the approach taken by this thesis. Finally, a few closing thoughts and observations are made in the *End Note*.

Chapter 2

Performance Indicators and some surrounding issues

If each part of a system, considered separately, is made to operate as efficiently as possible, then the system as a whole will not operate as effectively as possible (Ackoff)

Introduction

This chapter looks at what Performance Indicators (PIs) are, how they can be used, and some of the surrounding issues. Furthermore, it aims to set the scene for many of the issues discussed later in this thesis. The term *performance indicator*, which originates from the business sector, has gained widespread usage in education and throughout the public sector, over the last twenty or so years: although as Jowett and Rothwell (1988) point out, PIs have been extensively used before this by the health sector and local government.

There are many thousands of different PIs, however at the most fundamental level Fitz-Gibbon (1996) points out that they are simply pieces of information which can tell us something about a system or organisation[5];

> A performance indicator can be defined as an item of information collected at regular intervals to track the performance of a system.
>
> (p. 5)

Goldstein and Spiegelhalter (1996) make the point that PIs can be used to summarise a number of pieces of information;

> In its widest sense a performance indicator is a summary statistical measurement on an institution or system which is intended to be related the quality of its functioning
>
> (p. 385)

However, on their own PIs may well have little meaning, and as Willms (1992) emphasises there is a need to contextualise PIs:

> An 'indicator' is simply a statistic describing some feature of the schooling system associated with its performance over time...Like most statistics, an indicator derives its meaning from its trend over time, from its variation within a sample, or from comparison to some standard.
>
> (p. 1)

Although there are many different PIs, the OECD (1995) suggest that they are predominately associated with exam and test scores:

> What are grandly called 'performance indicators' often come down to test or examination scores. Other indicators of schools performance (the most popular being drop-out rates, truancy rates, and the destination of graduates from the system) have persistent methodological problems, and are even more likely than test scores simply to indicate the social background of pupils. Most other indicators are

---

[5] This definition is provided with a warning of the dangers of trying to define words with more words.

> really addressing questions of function, not performance, although they can be important in building up a picture of the school.
>
> (p. 46)

This would seem to be a slightly defeatist view, and perhaps misses some of the potential value of PIs: for example, in terms of the functioning and long term performance of schools and the overall education system.

Two uses: Accountability and Management

Jowett and Rothwell (1988) point out that PIs have two main functions; externally for accountability, and internally as a management tool. In terms of accountability, the changing nature of public sector management, during the 1980's in particular, set the scene for the citizen's charter (DFE 1994) and the 1988 education reform act (ERA 1988), which established the key role of PIs in holding schools accountable. From the point of view of management, schools have a long history of using PIs for schools improvement, at all levels from individual class teachers, to whole school policies (Willms 1992, Cohen 1981).

From the management perspective, Fitz-Gibbon (1996) argues that the key to an effective performance indicator system is the ability to provide a wide range of information to the relevant people, which will enable them to improve the process, ie. to provide feedback. In a similar vein Willms (1992) highlights the need to provide information to schools, as well as feeding forwards information to policy makers;

> Monitoring systems can provide useful feedback about the strengths and weaknesses of a schooling system. They can enable us to gauge whether inequities in academic achievement are increasing or waning. They can serve to assess whether a school or district reform is having a significant impact. they can help us learn why some schools are performing better than others, and thereby raise critical questions about educational policy.
>
> (p. x)

The use of PIs for internal management can be summarized with reference to Willms who states that PIs (as part of monitoring systems) should;

- Contribute to the working knowledge of both teachers and administrators
- Improve schooling
- Reduce inequalities
- Be used in conjunction with other data and observations
- Induce debate about policy and practice
- Inform the decision making processes

(Based on Willms 1992: 8)

For both Fitz-Gibbon (1996) and Willms (1992), performance indicators provide an important feedback mechanism to help manage education, which they point out is a *complex* system. Complexity exists not only in education, but many other areas of the public sector (Wilkinson 1997), as well as many businesses (Dent 1996). A key feature of complex organisations, is that they can not be managed effectively in a straightforward techno-rationalist manner. An understanding of the many component parts of the organisation and interaction between these is required, and the use of PIs, although essential, is not a clear-cut technical activity. Plesck and Wilson (2001) with reference to the health service provide a useful illustration of the evolution of a complex system approach;

> Management thinking[6] has viewed the organisation as a machine and believed that considering parts in isolation, specifying changes in detail, battling resistance to change, and reducing variation will lead to better performance. In contrast, complexity thinking suggests that relationships between parts are more important than the parts themselves, that minimum specifications yield more creativity than detailed plans. Treating organisations as complex adaptive systems allows a new and more productive management style to emerge...
>
> (p. 746)

Within education the inherent complexity means that, policies, interventions and even the indicators themselves may have unpredictable, unintended and sometimes chaotic consequences. Therefore, policies and PIs can not simply be used as 'leavers and dials', to produce highly predictable results. PIs have a more subtle and delicate function of providing feedback which is an essential ingredient in the learning process.

However, for government a complex or iterative approach to school management and improvement is not an attractive or even perhaps viable option, because of the 'modern' adversarial political climate[7]. This climate demands tough decisive action, with performance indicators being at the forefront of the 'zero tolerance' crusade on raising standards, and eliminating 'poor' performance. This perspective, with PIs as controls and policy instruments is evident in the OECD study, *Schools Under Scrutiny* (1995), which states that performance indicators are used;

- to generate data on national standards;
- to monitor the progress of reforms and to ensure they are put into practice;
- to evaluate the effectiveness of certain policies;
- to make sure that schools are complying with regulations;
- to monitor value-for-money;
- to improve the responsiveness of the system to the demands of society;
- to elicit information which would improve the quality of parental choice;
- as part of a system of accountability seen as integral to the democratic process;

---

[6] This is perhaps an over generalisation, some past management thinking has recognised complexity, see chapter 4

[7] Bennett *et al* (2000) point out that the omnipresent Ofsted inspection process, which is in effect an instrument of control, encourages a techno-rationalist approach to management, and consequently rejects the notion of complexity.

- to identify the strengths and weaknesses of individual schools as part of a national improvement strategy;
- to highlight schools with serious problems and attempt to address them;
- to assess the professional competence of teachers;
- to impose or encourage new or more effective ways of operating;
- to raise levels of pupil performance (at a national, local or individual level)
- to create 'learning organisations' - institutions which embody a culture of self-managed improvement and evaluation;

(p. 24)

The last point, the creation of 'learning organisations' seems at odds to the others, and illustrates some of the tensions which exist within the education system. The learning organisation approach recognises organisational complexities, as opposed to believing that policies and interventions can be precisely planned and implemented. This latter view, a techno-rationalist perspective, believes that PIs can reasonably precisely measure the specific parts of the system, whereas a complexity perspective recognises the limitations of PIs, and uses them more as 'hints and guides'.

The government's position on the use of performance indicators does appear a little inconsistent. On the one hand espoused policy and documents from the DfES, and its previous incarnations, (for example DfES 2001d – *Strategy to 2006*, DfEE 1998b – *Target setting in schools*, DFE 1994 – *parents charter*) advocate a techno-rationalist use of PIs, and without any recognition or warning of potential dangers. On the other hand, policy documents such as the white paper *modernising government* (Cabinet Office 1999) advocate techniques such as the 'balance scorecard'[8], which it is widely accepted (eg. Woodcock 1998) would help reduce possible dysfunctional effects of PIs. Furthermore, documents such as *FABRIC – A framework for performance information*, (Audit Office 2001) which was produced jointly by the Audit Commission, Cabinet Office, National Statistics, National Audit Office and HM Treasury, not only advocate a balanced approach to the use of performance indicators, but explicitly draws attention to the danger of PIs causing 'perverse incentives'.

Schools can use a wide variety of PIs (CIPFA 1988, Carter *et al* 1992), for example the document *Performance Indicators: An aide-memoire from DES* (Education 1989) identified 50 main groups of PIs, of which each has many individual specific indicators covering many aspects of the whole system. As would be expected this list includes objective indicators such as exam results and staff pupil ratios, as well as other more subjective indicators, such as the amount of graffiti in the school, and the extent of links with local businesses. These individual indicators can be considered in terms of how they are used, (eg. different types of monitoring), and which parts of the overall education process they measure.

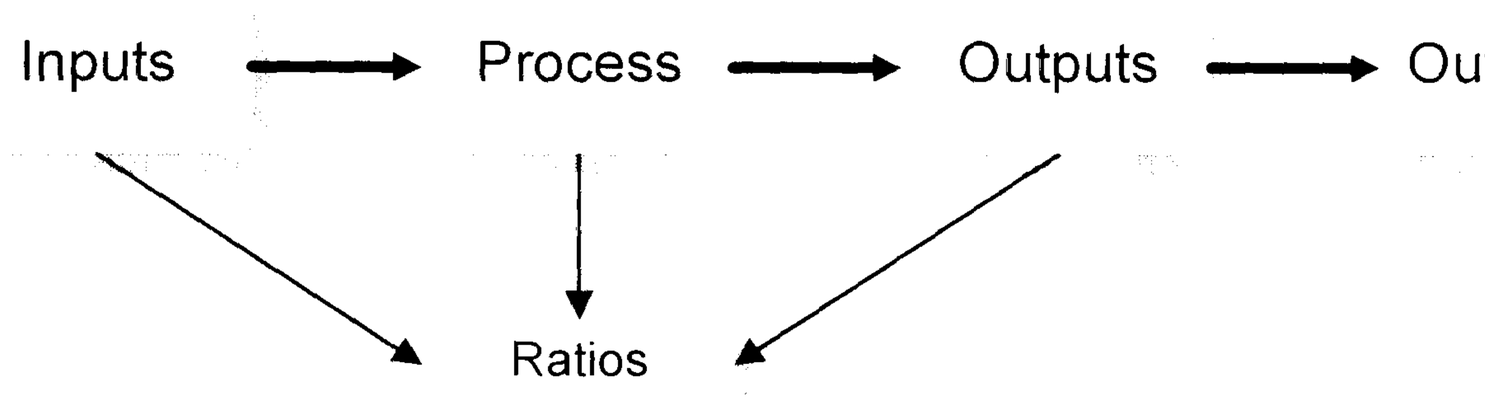
---

[8] In essence using a wide balanced range of indicator to measure overall performance

Individual indicators: Input to Outcome

In education, as with most other organisations, individual PIs can be considered to measure the performance at four different stages in the overall process (Fitz-Gibbon 1996). A fifth group based on ratios or a combination of indicators can also be formed. Although the model shows four discrete stages, in practice there is overlap and ambiguity between them. for example some outputs may be considered outcomes, and vice versa.

Figure 1: Indicators – Input to Outcomes



Input performance indicators are typically numeric and based on the resources available to the organisation. Examples include the number of teachers and support staff, finance such as the Age Related Pupil Unit (ARPU) and specific capital grants. Input indicators are frequently used by schools and governments for planning purposes, although they may also be used for political purposes, for example to argue that the service is improving because of an increase in spending or the number of teachers. In general, inputs are relatively objective, easy to measure, but say very little about the quality of the education process.

Process indicators look at how the actual processes, ie. teaching or educating are performing. In general they can be used on an 'as required' or continuous basis, rather than on a fixed basis as with external tests and exams. Within a school they would typically form part of the internal management system, and may well be used for diagnostic monitoring. Formal systems such as those based on norm-referenced tests; for example, Cognitive Ability Tests (CAT) and Middle Years Indicator System (MIDYIS) from the National Foundation for Educational Research (NFER) and the Durham Curriculum Evaluation and Management (CEM) centre may be used internally by school and teachers for management and planning purposes. Informal systems are also widely used, for example, at the classroom level this may be a simple tables test, or an awareness of how many hands go up. Informal systems represent much of the professional skill of teaching, indeed good

teaching is informed by a continuous stream of informal process indicators and nuances, particularly in the form of feedback.

Output indicators are used at the end of the process, typically when the child leaves the school, for example, National 5-14 and Key stage 2 tests. They may also be used to measure the effectiveness of a particular intervention at some point during the overall process, for example, the impact of the literacy initiative. Attendance and exclusion figures, which are reported on an annual basis, are also output indicators, however they may also be used as process indicators for the school to monitor its own performance, on a weekly, or termly basis. In a similar way, Key stage 3 tests which are taken part way through secondary school can be viewed as process indicators, however the decision to publish the results (DfES 2002) may make them output indicators.

Like input indicators, output indicators are attractive to policy makers because they are generally easy to collect, give an impression of objectivity, and are often available immediately after the particular process (Willms 1992). This allows them to be used to hold Local Education Authorities[9] and Education Authorities[10] - (L)EAs, schools and teachers accountable, and they may be used to justify actions against parts of the system which are deemed to be 'failing'.

Although output indicators can appear objective, in practice they can give a misleading or inaccurate indication of performance. For example, in schools with small cohorts there is a danger of substantial variation in the published results occurring simply by chance (Tymms 1993, Goldstein and Spiegelhalter 1996). Furthermore, output indicators often say very little about the quality of the process or effectiveness of the intervention, and may predominantly be an indication of other factors, such as prior attainment and/or soci-economic background. Output indicators can be improved by making allowances for factors such as these (for example an indication of the 'value added[11]'), and giving an indication of the likely error (for example, providing confidence intervals - see Fitz-Gibbon 1997, and Marshall and Spiegelhalter 1998). However, the government make the claim that giving output data in a raw format allows transparency, and argues that parents as consumers are quite capable of interpreting the data, and allowing for the necessary contextual factors (Tomlinson 2001).

Outcome indicators are measures of the longer term effectiveness of individual organisations, particular intervention or even the effectiveness of the entire system. They are by definition far more subjective than output indicators, and depending on what is measured, there can be greater

---

[9] England
[10] Scotland
[11] To allow for the different starting points of students

difficulties in making the actual assessments. Furthermore, they may well not be available at the end of the process, as it may be many years before the outcomes can be reasonably judged. However, outcome indicators frequently represent the most important aspects of the organisation, in that they can be a reflection of what it is aiming to achieve. In general, outcomes can be considered in terms of the *missions* and *aims*, where as outputs are *targets* and *objectives*.

In many cases there can be a degree of overlap between outputs and outcomes. For example, literacy skills which can easily be measured in numerical terms can be used as output indicators. Literacy is also an important outcome, however it is far more difficult to measure a long term love and appreciation of literature.

Fitz-Gibbon (1996) points out that it is important to differentiate between those outcomes which relate to the actual processes in the school, as opposed to those which are simply characteristics of the students or caused by other factors beyond the control of the school. However, in practice it can be very difficult to attribute responsibility, because in terms of outcomes many different teachers and schools may have contributed. For example, identifying 'who' taught a particular child to do percentages, or who inspired them to play for England.

Furthermore, many of the most important outcomes are both influenced by explicit government policy and the contribution of other organisations. For example, it is difficult to attribute responsibility for areas such as; physical and mental health, employability, artistic appreciation, crime, and teenage pregnancy. The concept of 'joined up' government, local and national, should in theory improve overall performance in many of these areas, but a new approach to performance measurement may be required. This is one of the great challenges for PI systems, and is discussed in chapter 14.

Ratios can be formed by combining different measures. They are very easy to produce, and can be very powerful and informative; however they can also be completely meaningless and dangerous. A number of ratios are routinely reported such as the pupil / teacher ratio, and the proportion of statemented children.

Ratios or ratio analysis is extensively used in business performance reporting. For example the ubiquitous Return On Investment (ROI), is based on the cost of the investment (input) and the value of return (output). Similar ratio's are being proposed for use in education. The Treasury is considering a ratio of 'value for money' (Mayo 2000) which would be based on the costs (inputs) and exam / test results (output). In secondary schools for example, this would give the 'cost per GCSE point', and this measure is already available in the independent sector (Sunday Times 2001).

Monitoring systems

Performance indicators play an important part in monitoring systems. Richards (1988) points out that there are three main types of monitoring; *Compliance, Diagnosis, Improvement*, and again there can be some degree of overlap between the different types.

Compliance monitoring, as the name suggests ensures that the organisation complies with specific conditions. These might be statutory requirements from the Health and Safety legislation, or items which are required by advisory documents, from bodies such as the DfES, Ofsted, and the Audit Commission. For example, an indicator from the *DES aide-memoire: Performance Indicators* (DES 1989) ensures that schools have a 'statement of objectives' (para. 41). In essence, compliance monitoring aims to ensure that schools have the necessary resources and other inputs to be able to perform satisfactorily (Willms 1992).

Diagnostic monitoring uses process indicators to assess or determine whether certain conditions and levels of performance are being achieved. In education, Willms (1992) points out that the goal of diagnostic monitoring is to determine whether specific aspects of the curriculum are being mastered by the majority of the pupils. Diagnostic monitoring is carried out frequently and provides feedback to both teachers and pupils. Although external tests may be used, this form of monitoring is 'owned' and controlled by the school. In general diagnostic monitoring can not be used to make valid comparisons between teachers and schools; however, in practice indicators (eg. YELLIS results) may be used in performance related pay (PRP) systems, such as threshold and performance management assessments (DfES 2001f).

Performance monitoring is primarily concerned with assessing how effective one organisation is compared to another. For example, businesses may be judged by profitability or earning per share, hospitals by waiting lists, and schools by GCSEs or SATs (Key Stage tests). In education the Key Stage tests are criterion referenced, and then the threshold are set to make the tests 'the same' as the previous years (see, Quinlan and Scharaschkin 1999), to allow valid comparisons to be made between individual schools and over time. There are however problems with this approach (Gold 2002, Wiliam 2002) both in terms of the testing procedures and the effect they have on schools. In an attempt to control for soci-economic differences between the schools, the official DfES PANDA (DfEE 2000b) benchmarking system uses the take up of free school meals as an indicator of social deprivation. However, as Croxford (2000) and Coe (2001), point out this is far from perfect, or statistically acceptable; indeed giving such spurious indications of fairness may in itself dysfunctional.

Data quality

Quantitative performance indicators or KPIs which are used for performance monitoring may be obtained and processed in a number of ways. How this is done will affect their quality, and *should* affect how they are used. Fitz-Gibbon (1996) identifies four kinds of data: *raw data, comparisons, fair comparisons and really fair comparisons.* These form a hierarchy, with raw data being the lowest quality and experimental the highest or 'gold standard'. She emphasises that an appreciation of the limitations of these different types will reduce possible misunderstandings or misinterpretation.

Raw Data as the name suggests is simply the measurements on their own with no comparisons or allowances being made for different circumstances. For example, individual GCSE scores are raw data, from which very little inference in terms of performance should be drawn. However, there are certain benchmarks, such as the 5 good GCSEs (A*-C) which are generally viewed as good[12], although they are something of a throw back to the old standard of 5 'O' levels, and indeed before that in the 19th century, the 5 pass matriculation for university. Therefore, to derive any reasonable understanding from raw data, comparisons have to be made (Willms 1992).

Comparative Data allows inferences to be made by processing raw data or comparing it with other data. There are a number of ways that this can be achieved, with at the one end of the quality continuum indicators being little better than raw data (eg. league tables), and at the other, modelled data which is of a quality approaching experimental data.

Rankings or league tables based on raw data can be used to provide crude comparisons of performance. They are available in the form of the official DfES / LEA performance tables, and the many media league tables, which generally use the same data, but sometimes with additional information. They are easy to create and provide a superficial indication of performance, however, as argued by many they do not provide an accurate or fair indication of performance (eg. Fitz-Gibbon 1996, Goldstein and Spiegelhalter 1996).

Benchmarking can improve the quality of the data. As previously mentioned the official method (PANDA / Autumn package (eg DfEE 2000b), make adjustments to the results based on FSM entitlement, or prior attainment based on results of previous KS tests. Benchmarking can also be carried out on a LEA or group of schools basis using other schools with similar characteristics. The

---

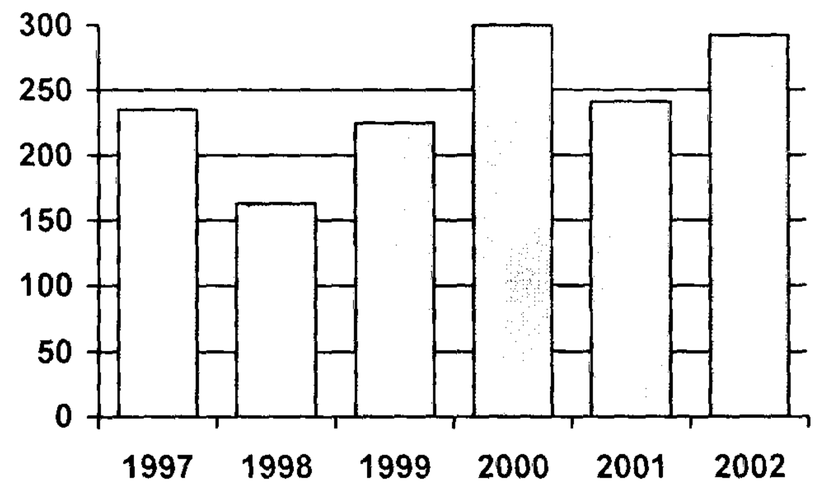
[12] Although of course for many children three grade 'E' would be good, or 10 'A's – it's very subjective

results from the these systems is not routinely made public, although they will be used in Ofsted inspections, which are publicly available[13].

Croxford and Cowie (1996) looked at the relative performance of secondary schools in Grampian. Not surprisingly there was substantial variation, with a difference of five and a half standard grades between 'best' and 'worst'. However, when allowances were made for the differences in intake, based on *pupil background, school social context, neighbourhood deprivation, stability and gender*, the 'difference' reduced to just one standard grade. Applying confidence intervals to these findings meant that only a few schools can be identified as significantly better, or worse, than 'average'.

Time scale comparisons can be used, for example to show an increase or decrease in average GCSE scores over a number of years. The official performance tables show a summary of each schools performance over the last four years as an *improvement measure*. However this can be quite misleading, for example;

Figure 2: Sample KS2 performance table results



Results based on the sum of percentages of pupils gaining level 4 KS2

This small[14] school was one of the Sunday Times top 100 primary schools (Sunday Times 2001) in 2000, having been below 'average' two years before. During this period the staffing and all of the other key variables remained constant. The only significant variable was the children taking the KS2 tests. It is pertinent to note that in 1998 when the results were 'low' the governors of the school successfully applied for the Head to be given early retirement on the grounds of 'ineffectiveness'; he left in 2000!

[13] www.ofsted.gov.uk
[14] Cohort 11-14

Furthermore, this yo-yoing in apparent performance is a feature of most, if not all, school KPI performance. The article, *League Tables ups and downs* (BBC 2002) reports that for 2002, 52% of schools improved and 47% got worse. Hardly a surprising finding, and in itself a good illustration of the perils of basing too much (or perhaps anything) on league tables. Furthermore, the article draws attention to some of the most spectacular rises, for example one schools scored 41 in the year 2000 and 240 in 2001; needless to say similarly spectacular falls must have also occurred.

Fair Comparisons or 'residuals' which are also known as 'Value Added' (Fitz-Gibbon 1996), aim to give an indication of the contribution made by the school over a particular time scale. This can be calculated by finding the difference between two test scores at different times, and comparing this improvement to the improvement made by others. This residual amount may be positive, ie. better than expected, or negative. Value added can be calculated at the pupil, class or school level, with allowances being made for prior attainment, which has been found to be the most useful predictor (Fitz-Gibbon 1997).

There are two main systems used by schools on a national basis, the DfES 'Autumn Package', and the Durham CEM centre systems (including PIPS, MIDYIS and YELLIS). The examples below are based on the secondary systems, although the underlying principles are very similar for primary schools.

The KS4 Autumn package (DFEE 2000b) uses the results from the KS3[15] tests to predict GCSE / GNVQ results. These are provided on a matched basis for Maths English and Science, and on an average point score for a range of other subjects. The results for each child can be plotted on the provided graphs, which show median and quartile lines. The residuals or value added can be read off for individual pupils, and these results can be summarized on a class or school basis.

YELLIS consists of a test containing two sections, vocabulary and mathematics, from which a combined score is obtained. This score is then used predict GCSE grades in a wide range of subjects. However, the regression coefficients vary considerably between subjects (Fitz-Gibbon and Vincent 1997), therefore individual regression equations are used for each subject. Coe (2001) reports the correlation between prediction and results for the five most popular subjects as being between 0.5 and 0.8. Again the results can be used on an individual pupil basis, as well as at the class and school level.

---

[15] Tables are also given for KS2 to KS4.

More sophisticated multi-level models can be used with other factors besides prior attainment (see for example Goldstein 1999 and Willms 1992). However, in practice this extra degree of sophistication does not add in any significant way to the accuracy. For example, Coe (2001) found a correlation of 0.99 between OLS[16] regression and a multi-level model, and Tymms (1997) found that the error for OLS were only slightly greater for smaller schools. Even given these possible errors OLS has a substantial advantage in terms of transparency and understanding, which as the Value Added National Project (Fitz-Gibbon 1997) pointed out, are an essential part in the meaningful interpretation and use of performance data.

Currently, value added data is used internally by schools for their own management. However, the government are committed to publishing some form of value added data alongside the KPIs. Although value added figures do give a fairer indication of performance there are still potential problems in their publication, as Goldstein (1999) warns;

> The results [value added indicators] are in the nature of screening instruments that can indicate where problems may be present, but which are not precise diagnoses. In particular their use as public accountability measures, e.g. in the form of performance tables or 'value added league tables' is inappropriate and would destroy their credibility and usefulness. If they were ever to become 'high stakes' pieces of information like the current DfEE league tables of examination results, then they would inevitably become distorted and no longer reflect any underlying reality of school performance. The same conclusions, of course, apply to any future national scheme which attempts to derive value added measures.
>
> (p. 21)

In the United States, Dorn (1998) pointed out that shortly after the publication of data from the Tennessee Value Added Project, the media re-published this data in the form of league tables. This may in itself encourage behaviour to maximise the value added score, which may not be in the best interests of the school or children. There is widespread support for value added systems, many schools have found them extremely useful, but they must still be used with care.

Really fair comparisons allow the effect of different contextual factors to be controlled for by carrying out some form of experiment: ideally randomised control trials (RCTs). These are used extensively in medical trials, as well as in education for the development of evidence based practices, for example, evaluating different teaching methods. However, in terms of performance monitoring based on KPIs there would be a number of practical difficulties. To fairly judge a school's performance, many of the significant variables would have to be randomly assigned, for example the children, the building, the local area, and then the experiment would have to be repeated a number of times. Although this may well produce interesting results and question many of the underlying assumptions in education, such a proposition is impracticable on a large scale;

---

[16] Ordinary Least Square – used by YELLIS

therefore we are often limited to comparative methods when comparing results. (This issue is discussed further in chapter 8, *A qualitative or quantitative approach*)

Proxy Indicators

When assessing the performance of most organisations including schools there are many indicators which could be used, however in practice only a limited number can be used effectively. Rangone (1997) points out that people generally have difficulty comprehending and summarising large numbers of performance indicators, and therefore it is necessary to provide a number of key indicators. These key indicators, or KPIs are used as proxy indicators of overall performance.

In general proxy indicators, with an understanding of their limitations, are very useful management tools. Boland and Fowler (2000) highlight some of the potential problems and limitations of their use in the public sector, in particular that many of the most important issues may not even be measurable[17]. Brignall and Modell (2000) point out specific proxy indicators are often chosen simply because the data is readily available or can be easily collected, rather than being specifically designed for the purpose.

A number of examples of proxy indicators have been previously mentioned, for example the uptake of free school meals, as an indicator of deprivation. Many researchers have pointed to the limitations of this indicator (eg. Croxford 2000, Plewis and Goldstein 1998). Proxy indicators can only tell a part of the story, indeed for many a 'leap of faith' is required in their interpretation, and they may still give a misleading impression or miss the point altogether.

High Stakes Indicators

The majority of educational performance indicators are relatively low stakes, such as those used by schools for internal management, or statistical returns. However, KPIs are by definition, high stakes. The *American Educational Research Association* describe high stakes indicators as those which "carry serious consequences for students or educators" (AERA 2000: 1).

In general high stakes indicators are made public, indeed the publicity may increase the stakes, and they are often associated with the principle of public accountability. For some indicators most people, including those whose performance is measured, will support publication; for example indicators relating to industrial safety. The key issues are relatively straightforward, and the indicator may well discourage bad and unsafe practices, both by the employer and employees.

---

[17] Gorz (1989) goes further, and suggests in some circumstances, the performance of caring professions, such as teaching, can be inversely proportional to the quantifiable outputs (p. 143)

However, for other high stakes indicators, particularly those associated with complex systems, the underlying performance issues are far less clear, and there can be problems in making these indicators public. For example, William Sanders who designed the Tennessee value added system, which was one of the first and most prominent high stakes indicator system, points out that the system was never intended to be made public (Marcus 2000). In England schools are not permitted to publish results from the CEM centre value added systems, nevertheless the results from these test can be high stakes, for example, in terms of teacher appraisal.

High stakes indicators are linked to the often quoted government mantra of, "driving up standards" in the public sector. In practice this can take the form of combining pressure and support, for example Barber (2000b) describes the government's approach to education as "High challenge, High Support" (p. 5), with the challenge being set in terms of high stakes proxy indicators (KPIs). This view is reflected in a leader in the *Times Educational Supplement* (TES 1999) which points out that we are out of step with the rest of Europe in terms of the wide spread use of high stakes testing; however, it does argue that this is necessary to overcome years of under achievement in Britain by working-class and ethnic-minority children.

As well as KPIs being high stakes indicators for children, schools and LEAs, they are also 'high stakes' for the government, both collectively and for individual ministers. David Blunkett staked his career as Education secretary on the primary KS2 targets for 2002 (Smith 1998). Since making this promise he has changed jobs, and his successor dutifully took on this promise. In a less public way ministers are held personally responsible by the treasury for meeting 'their' Public Service Agreement (PSA) targets[18] (Treasury 2000). Those ministers not meeting their targets have to appear before the Chancellor's 'star chamber' (PSX committee) to be cross-examined (Sherman 2001).

Reliability and Validity

The issues of reliability and validity are of fundamental importance to the design and interpretation of performance indicators (Kerlinger 1986). Reliability can be defined as the level of consistency shown by an indicator system; a highly reliable examination system will consistently give the same score to an individual student, or in the case of a group of students their rankings will remain consistent. Furthermore, these results should remain constant over time, with different examiners and types of questions.

---

[18] KPI targets (GCSE and KS2) are derived from the PSAs.

Shepard (2000) has highlighted the problem in high stakes testing of transferring knowledge and skills from one test to another. Tests in different formats to the high stakes test may give different results, even though the same skills and knowledge are assessed, thus suggesting that the tests might be unreliable. However, this may simply be due to students being extensively and exclusively prepared for the one type of test[19]. In terms of KPIs although there have in the initial stages been quite spectacular rises, other tests, for example the PIPS reading test (Durham CEM) have not reflected this rise (Tymms and Fitz-Gibbon 2001).

The size of the cohort will have a bearing on the reliability of a test. For example, more accurate inferences in terms of school performance can be drawn from a hundred children doing a test, rather then ten[20]. This can be improved further by combining the results from successive year groups. Currently, KS2 results are not published for schools with less than 11 children taking the tests, because of the inherent statistical unreliability of small groups. Although the results will be used (and potentially statistically mis-used) by others such as the Governors, LEA and Ofsted.

With respect to validity, a performance indicator can be said to be valid, if it measures what it claims to measure. However, the main problem in the social sciences is knowing what to measure (Kerlinger 1986), and this takes us back to using proxy measures, with the attendant shortcomings. Furthermore, PIs may be valid at one level of the system, but not necessarily at another, Wiliam (1994) describes this with reference to KS assessments:

> When we make assessments, we are hardly ever interested primarily in the actual items we are testing – we are interested in the ability of the assessment results to 'speak' for other aspects not tested. Assessment information is useful because we believe that we can generalise from the actual items assessed to some wider domain....
>
> (p 17)

This situation can be illustrated with the example of a KS mathematics test question. The question may well give a very valid indication of a child's ability to calculate the area of a triangle (at that particular moment). The test as a whole may also give quite a valid indication of the child's mathematical ability. However, it may well not be valid to draw inferences about the mathematics teaching in the school, the child's overall academic ability, the quality of the particular teacher, how good the school is, and so on. This view is supported by Stobart (2001), when using the *Eight stage threats to validity* model (Crooks *et al* 1996), argues that using KS tests to measure year on year changes in performance is particularly questionable.

---

[19] Davis (1999) points out that 'teaching to the test' is not a new problem. He also suggests that modern examination methods should be able to avoid these problems.
[20] See also Tymms (1993) who looks at the effect of natural variation on test results, and Fitz-Gibbon (1997) with reference to confidence intervals.

Indicators such as eligibility for FSM may well be reliable, and valid, in that they consistently measure what they claim to measure. However, once they are combined with other indicators such as KS tests to give a ratio or a contextualised performance indicator, the resultant figure may no longer be valid for a particular purpose, such as measuring the relative efficiency of a school. (see for example Croxford 2000 earlier in this chapter).

Although validity and reliability have been considered separately, in practice they are closely related and interdependent; and as Davis (1999) points out, increasing one aspect (eg reliability) may cause a corresponding reduction in the other[21]. Furthermore, Stobart (2001) who states that "Validity is not what it used to be" (p. 27), argues that when considering validity other consequences of the assessment systems should also be considered[22]. In effect validity can not be considered in isolation, but rather it should be considered within the overall context of the organisation[23].

The broader management context

To more fully appreciate the potential effects of PIs it is necessary to consider the broader context within which they exist. As Pollitt (1993) points out PIs are a part of, and interact with, many other aspects of the organisation and management of schools:

> Particular performance indicators (bed turnover in hospital for example, or staff/student ratios in education) appear by themselves to be relatively neutral, technical artefacts. Yet to appreciate the full significance one also has to take account of the other techniques which accompany them, the model of management within which they are developed and the broad assumptions underlying both.
>
> (p. 12)

Looking at the actual management processes of which PIs are a part, Child (1969) pointed out, (some thirty years ago), that this can not be thought of in terms of simple techno-rationalist activities, but rather it is necessary to consider the social and authoritative aspects;

> For the development of management thought, and the awareness of the management role which it reflects, cannot be understood merely through reference to the technical aspects of managing. This is because, in addition to its technical function, management is a system of authority through which policy is translated into the execution of tasks; and. . . an elite social grouping which acts as an economic resource and maintains the associated system of authority.
>
> (p. 13)

---

[21] He cites an example of increasing the reliability of an interviewing process by structuring and formalising the questions for all of the interviewees, however this may reduce the validity, for the interviewees by not allowing individualised questions and responses.

[22] Stobart argues that the responsibility for this lies with the designers of the tests, ie the Government though the Qualifications and Curriculum Authority (QCA).

[23] Implicit is this view is that issues relating to reliability are considered alongside or as part of validity.

Chapter 4 develops these points and arguments by using *organisational theory* to look at the broader contextual issues of PIs.

Levels of management

In schools, like any other organisations, it should be reasonably clear who or what body, is responsible for the various management functions. However, in education this may not always be the case; for example, are schools managed and controlled by the Head and staff, the Governors and parents, the local authority (L)EA or Ofsted / HMI and the government? This question cannot be easily or simply answered, as each of these groups has some role, and even this may vary depending on the measured performance of the school.

The underlying principle of 'local management of schools'[24] (DES 1988) would seem to suggest that the key management decisions should be taken locally, i.e. away from central government. Indeed, at the strategic level, the Teacher Training Agency (TTA), state that Heads are responsible for the, 'Strategic direction and development of the school', (TTA 1998: 9). The Secretary of State for education (Morris 2001a) claims that the current government wants to encourage schools to have more freedom and autonomy;

> We want to free the energies, talents and creativity of heads, governors and teachers to support them to achieve higher standards and to enable them to innovate and move towards earned autonomy. Our schools already have higher levels of autonomy in decision-making than other schools in Europe[25].
>
> (p. 6)

Although there are claims by the government that schools do now have more freedom, in practice most schools toe the government line. For example, few schools would 'dare' to abandon the national literacy strategy, even if they were confident that it was not in the best interest of their children, and they had themselves a better strategy. Furthermore, Morris's view does not appear to concur with that of her predecessor. David Blunkett justified the increased central control by government;

> ... because our experience led us to understand that central government carried responsibility without power. It would be to delude the electorate to suggest the Government could deliver change without the mechanisms to deliver it. We are taking the necessary structural powers to ensure the people on the ground do their job.
>
> (Carvel 1998)

---

[24] Strictly speaking LMS refers to financial aspects of school management.
[25] Unfortunately this document (speech to the Social Market Foundation) gives no details of the basis of this surprising claim.

Blunkett's view may well be more familiar to those associated with the education system. And it is the view shared by many other writers (eg. Baker 1994, Pollitt 1993, Whitty *et al* 1998); for example Glatter (1999) points out that;

> The [Education] system was until the 1980s widely regarded as one of the most decentralized in the world. Today it is among the most centralized of the advanced industrial countries. The trend appears if anything to have intensified since the coming to office of the Labour government in May 1997.
>
> (p. 254).

In comparison to other countries, there appears to be little evidence to support the view that heads have a higher degree of autonomy than their international counterparts. Mahony and Moos (1998) contrast the increasingly managerialist approach to education in England, with the more democratic and inclusive approach in Denmark. Although, Power *et al* (1997) found that an increase in central control was a feature of the five countries they considered[26]; they cited research by Granström (1996) which suggested that, although claims were made that the educational reforms would bring schools greater participation in policy making, in reality there was less. Indeed some in government take a more extreme view, Tim Brighouse an adviser, stated that; "Our national curriculum is more nationally prescriptive than any other state and is more so than the Stalinist regimes of the USSR" (BBC 2002). In contrast to these observations, Ball (1998) points out that the increased control over education has been accompanied by correspondingly less governmental control of commercial organisations.

The changes in the political processes, in the form of a 'New Public Management' are considered in more detail in chapter 5. However, at this stage Broadfoot (1999) usefully illustrates how the element of control in the system has changed, resulting in her claim, that the English education system is one of the most 'controlled' in the world;

> Only a few years ago, educational visitors from other countries who visited England found it almost impossible to credit that there was no national curriculum, indeed very little central direction of any kind to constrain the activities of either Local Authorities or schools. They marvelled at how free individual teachers appeared to be in their own classrooms to decide what to teach, how and when...Such an experience would not have prepared them, however, for a similar visit today when they would find an education system that is now arguably not only as tightly controlled and centrally directed as any in the world...
>
> (p. 2)

In practice control is in part exercised through a seemingly never-ending stream of new policies. Bassey (2001) points out that in the first two years of government, David Blunkett issued an unprecedented 332 directives to schools and launched 52 centrally managed initiatives. In such a climate of continual change, and strong directions, it would seem difficult for schools and teachers

---

[26] England and Wales, USA, Sweden, Australia and New Zealand.

to develop their own innovative strategies. For most it is far easier to take the safe 'pre-packaged' options, for example the QCA schemes of work – indeed this option may well allow more time for teaching. With respect to LEAs, Meredith (2000) points to their reduced opportunity to contribute to the strategic development of education, with many directives, and the requirement to produce numerous plans which have to be 'approved' by the DfES.

Moving away from strategic planning, to the day-to-day management (or the operational level), the LMS principle and the 'new educational management' have brought far more decision making down to the school level. For example, in the case of personnel management, most schools and their staff would probably agree that this is best dealt with at the school level, rather than by the LEA. Likewise many other functions such as cleaning, meals, purchasing and commissioning training are also probably best organised at the school level. Farrell and Morris (1999) found that both secondary, and perhaps surprisingly, primary schools were generally in favour of the increased responsibilities that accompanied LMS. However, evidence from the NAHT (2000) found that heads were increasingly having to do more 'low-value' tasks, at the expense of high-value tasks, which they felt would benefit children's education; for example, making decisions about who cuts the grass, has little impact on the quality of education.

LEAs and Unions

Some thirty years ago, Chitty (1989) suggested that top-level education strategy formulation, was essentially a partnership between the DES, LEAs and the unions. The reforms of the 1980's and 90's have almost entirely sidelined the local authorities and unions, with the claim that schools should be, and indeed are now more responsive to parents (Barber 2000b); whether this is in fact the case is very much open to question, and is considered later. Nowadays the position of the LEAs is quite confusing, and perhaps cynically their role appears to be there to 'take the blame'. They still have some strategic responsibilities for the provision of school places, 'education other than at school', target setting, and some aspects of special needs support. However, their role now is more of a 'principle contractor', who awards and oversees 'sub-contractors' who provide educational services. Furthermore, there are an increasing number of LEAs which no longer exist, and have had their functions handed over to private sector contractors.

The position of the teaching unions is also unclear. Although the term unions is used here collectively to refer to the main teaching unions and associations, as well as the two Head's associations, it should be noted that these separate bodies have different relationships with the government. In terms of strategy their role appears to be rather limited, with the newly formed General Teaching Council (GTC) being recognised by the government as representing the views of teachers. Furthermore, the DfES strategic plan *Delivering Results: A Strategy to 2006* (DfES

2001d) does not include the teaching unions as one of their 'partners', whereas others such as the GTC and private sector training providers are.

The Government

Although the collective term 'government' is used throughout the thesis, it is not a single coherent body; there are many parts of the government which contribute to the strategy, management and organisation of the education system. From the strategic perspective, two departments (besides the DfES) are particularly influential. The Treasury has responsibility for funding as well as setting the PSA targets from which KPIs are derived, and the cabinet office has an interest in ensuring that the targets are met, co-ordinating cross departmental policies, such as inclusion, and also the 'party political' aspects of performance and its reporting. Many other departments such as the Department of Trade and Industry (DTI), and the Department for Work and Pensions (DWP) and the Department for Transport, Local Government and the Regions (DTLR) also have an influence on the strategic direction of education, as well as other bodies such as Ofsted and the Audit Commission[27]. Other groups, such as the Social Market Foundation, DEMOS, the various religious organisations, the Confederation of British Industry (CBI) and many of the various 'edu-businesses', as well as the unions have some degree of 'back door' influence.

In reality the education system is very much controlled by the government and its associated bodies, with specific operational functions being passed down to schools. Although the Government has called on heads to be more creative and entrepreneurial, in effect they are more akin to 'branch managers', rather than free thinking chief executives or DotCom entrepreneurs. Likewise, LEAs have functional responsibilities, rather than any significant managerial input.

Conclusion

This chapter has introduced performance indicators and some of the surrounding issues, as well as how, and by whom they are used in the school management process. The first part looked at how PIs may be used as management tools and / or accountability systems. This is an important distinction, because the former may well allow the users to use and control the indicators as they see appropriate, where as the latter very much presumes that the 'users' or schools have to be controlled, and the PIs are an important part in this process. These differences in use and approach will be considered throughout the thesis. The other particularly important issue which was discussed is how PIs fit in to the organisational structure, and from that how the organisational context will affect both how they are used and their effect on the organisation. This will also be considered further throughout the thesis.

---

[27] See DfES (2001d) for a list of some of their strategic partners

Chapter 3

The use of theory and experiences from other sectors

Reorganisation is a splendid method of producing the illusion of progress whilst creating confusion, inefficiency and demoralisation (Petronius Arbiter)

Introduction

Businesses have a long experience of using PIs as part of their management systems, and this chapter seeks to justify the use of *Organisational Theory* (OT), and *Management Accounting Theory* (MAT), in developing a theoretical framework. Some of the concerns from educationalists towards 'business theory' are firstly considered. Arguments that schools and businesses face similar challenges in measuring performance are made, and the emerging field of Educational Management Studies (EMS) and its convergence with generic OT is discussed[28].

Historically, there has been a degree of mistrust of businesses by a significant section of the educational community. The reasons for this appear to be largely ideological, although as Fielding (1999) points out, 'business language' has become linked to the 'crusade' for change[29], which many find uncomfortable. In practice, some of the concerns centre on the belief that commercial or market pressures have a corrupting influence on more venerated educational values, and that a consequence of making a profit is fewer resources for the children. In spite of these concerns the trend over the last twenty or so years has been towards a 'market approach', with the use of 'market type mechanisms', and an increasing role for the private sector[30] in the planning and delivery of education.

Ranson and Stewart (1994) point out that it was the Conservative Government's aim to make the public sector more business-like;

> Performance in the public sector can only be improved, the Government claims, by making public organisations and their management look as much like the private sector as possible... Indeed, the accelerating introduction of market-like competition to organise the provision of public services is intended to dissolve the boundary between public and private provision.
>
> (p. 26)

In the second term of the New Labour government there is continuing or strengthening support for private sector involvement;

> At the Confederation of British Industry's annual conference Education and Skills Secretary Estelle Morris today praised the management and financial skills of private companies who are working with the public sector to provide better schools.

---

[28] See Chapter 4 introduction for a discussion on the social sciences and scientific use of 'Theory'

[29] He makes specific reference to 'business' language of Barber and Sebba from the Standards and Effectiveness unit, and questions the need to "borrow the disfiguring, deeply dull language of performativity" (p176).

[30] The term private sector is used in it broadest sense, and as well as commercial companies can include, non-profit organisation, charities, and even other schools.

> "In the longer term I want the relationship between government and business in education to be based on a realistic appreciation of the strengths of each party, mutual understanding and respect. We need the structures to build real collaboration rather than being merely bureaucratic window dressing and help for each party in developing the skills and understanding to deal with the other, and we want a new Sector Skills Councils to drive this forward. They will provide a strong voice for business and we will listen."
>
> (DfES 2001g)

It was during the 1980s and 90s in particular that the involvement of commercial businesses in education increased. Although businesses have always provided some services, the relationship is now very different, with an ever increasing role and consequential responsibility, for the 'private sector'. Although in reality there would still seem to be much resistance and scepticism from many in the education world towards business involvement, (the unions for example) the government appears to make the rather surprising claim that these views no longer prevail. Estelle Morris (2001a), the previous Education Secretary, stated that;

> ...the deep suspicion between teachers and the business world in the 1970s and 80s was a thing of the past[31].

It is however suggested here that there is still significant resistance to business organisations, and as a consequence towards business theories and practice. Whilst the case is made for the use of generic business theory, this does not extend to arguing in favour or against the increasing privatisation or 'market' approach to education.

Some of the concerns

The basis of many of the concerns is the view that business theory does not recognise the complexities and subtleties of the education process. Bush (1999) points out that this goes back to the 'early days' of educational management training in the 1970's. And that much of the initial training was in the form of courses, offered in particular by the Open University, which drew heavily on industrial and American theoretical and conceptual frameworks. He goes on to illustrate some of the shortcomings of such frameworks and the significant amount of caution and "downright hostility" (p. 239) from some academics, who felt that business theory could not, and should not, be applied to education;

> To think about relationships within a school in business terms is to run the risk that people will be encouraged to behave in ways that are antithetical to certain fundamental educational values. The business analogy . . . tends to encourage a 'them' and 'us' relationship between head, senior staff, teachers and students, instead of stressing shared values within an academic community.
>
> (Taylor 1976: 41, Cited in Bush 1999: 239)

---

[31] Unfortunately no evidence was provided to support this view

For many this view continues to the present day. For example, Bottery (1999) refers to the global forces of which business management is part; "....have damaging effects upon richer, more human, conceptions of educational management." (p. 300)

The experiences of other public sector organisations does give some rise to concern. There have on the face of it, been a number of successful 'privatisations', however in general these have not been for complex organisations, but rather for those which deal with more simplistic functional tasks, such as the Teachers Pension Agency, or Driving Licence Vehicle Agency. There have also been a number of high profile failures, probably the most obvious being Railtrack (this is discussed in chapter 7).

Although many businesses do provide advice and resources for schools, particularly at the local level, this may well be of direct benefit to the schools and local community. However, there are also examples of where the motives are less philanthropic, and more commercially orientated. Whilst this is not necessarily bad or irregular, the education community, quite rightly, has to be careful and be aware of possible underlying motives. The potential to make money from education is increasing, with for example, foundation schools and various advisory and consultancy services. Aikenhead (1998) discusses the 'strings attached' to the free advice from the large accountancy firms. At times this advice may be more subtle but perhaps equally partisan; for example, John Mayo of Marconi PLC wrote a report; *Effective reporting in education*, for the Public Service Productivity Panel (Mayo 2000), which included a heavy 'sales pitch' for broadband technology for schools, which his company were hoping to provide[32].

In higher education too, concerns have been expressed about the increasing influence of the commercial sector (Ribbins 1999), with for example philanthropic sponsorship being replaced by a more targeted approach to research which will produce tangible returns. For example, Monbiot (2000) points out that he cannot find any example of sponsored research in which the sponsoring company does not have a direct interest. In a similar way Fitz-Gibbon (1996) questions the independence of Universities due to commercial pressures, and as she points out that this is not just an issue for the 'hard' sciences;

> "It must be the case that social science is more easily distorted by opinion and selective reporting and sheer wishful thinking, than is a 'hard' science."
>
> (p. 216).

---

[32] As it turned out this was a last ditch attempt at saving the company, and his job

New initiatives such as the city academies in this country (Budge 2000), which have a high degree of private sector involvement, and the Public Finance Initiative (PFI) are quite rightly viewed with a degree of caution (Pollock *et al* 2002, Monbiot 2002).

Measuring performance – similar challenges?

The ability to measure performance is vital to just about any organisation, whether private or public. For some organisations, measuring performance is relatively straightforward, however, for many complex organisations, including schools, it is difficult to do in a useful and meaningful way. Drucker (1990), points to the difficulties for non-profit organisations in that they do not have a 'bottom line' to measure; whereas profit organisations do; and it follows that this can be used as a measure of performance[33]. Whilst a measure of profit might be easy to produce, whether it gives an indication of anything meaningful is another matter. In a similar vein educationalists such as Bush (1999) and Glatter (1999) argue that education is fundamentally different from business because of the difficulties of measuring meaningful outputs, ie. no simple indication of profit. It is however argued here that measuring either educational or business performance is both complex and fraught with dangers.

Whilst businesses may be attempting to measure financial outcomes, and schools, educational outcomes; both types of outcome are the result of human effort, ingenuity and creativity. Ultimately, it is the contribution of factors such as these which all accounting systems are attempting to capture. Schools now most certainly do have a bottom line, which like business profit is measured with high stakes proxy indicators; only the units or terminology is different. The number of GCSEs or level 4 SATs is as important to a school as profit is to a business; and for both these figure may well say nothing about the actual performance, let alone future performance. Many different circumstances, unconnected with current or actual performance, can substantially influence a company's profit, or a school's performance. For example, expenditure on research and development will reduce profit, but potentially increase long-term performance or effectiveness, and a child missing the SATs exams through illness may well have a significant effect on the overall indicated year end results.

The reforms in the Health service provide some useful lessons for education. There are many intangible aspects of performance which are difficult to quantify, and which are often perceived by those working in the service, as well as those receiving the service, as being the most important; for example, friendly, caring and knowledgeable staff. Like education, the health service has gone

---

[33] However, profit as an indicator is far more open to manipulation then exam or KS test performance: eg. Enron (Shaw 2002)

through many substantial changes over the last twenty or so years. In general these were based on similar ideological arguments, such as the importance of quasi markets, and a general belief in the value of 'private sector management techniques'. This has led to the introduction of trust status and an 'internal market'[34]. As with education many of the reforms were, given the size and complexity of the organisation, hurriedly introduced and imposed. Jones and Dewing (1996) suggest that much could have been gained from first looking at the existing theory:

> Many of the issues ....are discussed in elementary textbooks on management accounting. It is a cause for concern that so few of them seem to have been put into practice in the early stages of the adoption of the internal market and trust status in the NHS. Many of the unintended consequences of change might have been avoided with improved forethought. Remedying the problems and getting users to adopt revised attitudes may take a long time.
>
> (p. 69)

They argue that the changes to the health service would have been far more effective had explicit reference been made to the well-established field of management accounting, and the same arguments are made in this thesis with respect to the educational reforms. However, it is not being suggested that the field of management accounting has all the answers, far from it. For example, in one of the classic texts, *Relevance Lost: The Rise and Fall of Management Accounting*, Johnson and Kaplan (1991) argue that;

> Today's management accounting information, driven by procedures and cycle of the organization's financial reporting system, is too late, too aggregated, and too distorted to be relevant for managers' planning and control decisions...And despite the considerable resources..... the figures do not measure the actual increase or decrease in value that has occurred during the period.
>
> (p. 1)

As with educational PI systems, business accounting has two relatively distinct roles: external reporting and internal management information. Companies are required by law to produce annual accounts, which are made public. These accounts are produced in a legally defined format, and like school governors' reports may also contain a variety of subjective and contextual information. The information is by definition historic, and does not necessarily give an accurate indication of the value added. In practice they are used to promote the company, and the key figure, profit, is used to determine how much tax is paid. Therefore, unlike educational reporting there may be an incentive in understating this; although bonuses and performance related pay may be based on this figure. In both situations, comparing two or more years (or accounting periods) *may* give a more accurate picture. And like education the information for some industries (eg Insurance) may be reported in league tables.

---

[34] The current government has reversed some of the principles, although the underlying 'market' philosophy remains.

Fundamental to any valid accounting system (business, education or whatever) is its ability to compare like with like. As mentioned in chapter 2, this may be in terms of comparisons with other organisations, within the organisation itself, or over particular time-scales. To do this there is a need to adopt a 'common currency' on which to base these comparisons. For commercial businesses many of the performance measurements can be converted or reduced to financial values. For example, the effectiveness of a marketing campaign can be assessed in terms of the increase in the value of sales, although this may not tell the whole story. Furthermore, as with education, many of the most important indicators, such as those that give an indication of quality and goodwill, and which may be an indication of future performance, cannot effectively be described in financial measures.

At one level, calculating a business's performance may appear straightforward; take for example the case of a company which buys raw materials, pays its workers and sells its products. The various costs can be added together, and compared to the income, to find the profit. This figure can easily be adjusted or refined to allow for other factors such as the depreciation of the machinery. Throughout the process, a common currency, money, has been used, and a clear unambiguous indication of the performance has been found - the profit. However, this does not allow for what may well be some of the most important factors of the business and its future; the value which has been added to the company. For example, new or satisfied customers (goodwill), and the ability of the employees to develop new products, (skills). Allowing for these factors in the accounting process is far more difficult, but very important for the decision making and planning processes. The same arguments can be applied to education: neat verifiable and reliable indicators such as the KPIs do give clear unambiguous figures, but like their business equivalents, they may well say very little about the value added and the skills of the teachers, and from this, the future performance.

One of an accountant's biggest problems is that of matching data, or allocating costs to income. Some costs can be easily allocated, such as materials and direct labour, however, others are far more difficult or impossible. For example, corporate promotions, research and development, and even the chairman's Bentley; what proportion should fairly be allocated to which products? The equivalent 'allocation' problem in education is attempting to trace who or what added which particular value to the process. For example, whether a particular child learnt to draw graphs in Maths, Science or Geography, or how should a school account for, and cost a teacher's time with a 'special needs' child.

Although the term profit is not used in educational performance measurement[35] conceptually there are similarities with 'value added'. Both profit and value added are calculated by measuring the inputs and comparing them to the outputs. In 'business' it is possible to have a negative profit, ie a loss, in the same way as it is possible to have a negative value added score. In both cases, the profit or the total value added score is an amalgamation of other figures, and on their own are of limited use. To manage effectively either types of organisation it is necessary to be able to 'drill down' to the underlying components.

Best Practice, Fads and Fashions

A common feature of both business and educational management theory is the continual search for the 'holy grail' of *the* best way. Many management theories and practices have been developed over the years, and one common feature is the claim that they were, in their time, *the* Best Practice. From this they tend to go into decline, and are then replaced by another theory, although sometimes they re-appear under a new label. The ephemeral nature of such theories has not reduced the government's desire to identify this mythical 'Best Practice' for use in the education system (DfEE 1999, TTA 1998), and more generally in the public sector (Cabinet Office 2000a, Smith 2000).

Within the business world, at any one time, there are many different interpretations as to what exactly is *the* 'best practice'. From the 'best' hundred companies in the UK, would come many different reasons for their success, and many different approaches to their management. For some this might be a highly structured and mechanistic approach, and for other the complete opposite. Furthermore, factors such as the diversity of their respective industries and the values and experiences of the owners would have an effect. Indeed the idiosyncrasies, or just simply style, of the owner may be the very reason for success. Businesses have the advantage over education in that they are able to learn 'best practice' from many other businesses and different sources, rather than just the one official government sources (NCLS (2002) framework, and the Hay McBer (2000) management model).

Currently, the Government (Cabinet Office 1999) is promoting the European Foundation for Quality Management (EFQM) excellence model, and using this as a definition of best practice for public sector organisations. This TQM type model[36] was originally developed in the business sector, and has become quite popular in some areas of the public sector (Cabinet Office 2000b). Bush (1999) however, points to the difficulties in defining the model of 'Best Practice' and indeed

---

[35] Although as previously mentioned 'for-profit' companies are an increasingly important part of the education service infrastructure.
[36] Total Quality Management - see chapter 4.

identifying who should decide what it is. Glatter (1999) suggest that one of the current education policy makers (ie the TTA) have identified 'Best practice' as being a hard 'techno-rationalist' approach, and therefore quite different to the EFQM model. In the business sector, the current fashionable 'Best Practice', is based on softer, i.e. more human orientated principles. Although, as Caulkin (1996) points out, the fashions in business are inclined to alternate every few years between hard and soft.

Many writers (eg. Fitz 1999, Phillips *et al* 1999) have highlighted this 'faddy' quality of business management. This results in a new business system or fashion appearing every few years, which is miraculously going to solve all of an organisation's problems, and increase profitability; or so the consultants claim (Shapiro 1998). The danger as Fitz (1999) argues, is that by the time these new systems are used in education, they may already be obsolete, having lost their appeal and credibility in the business world, 'fad lag'.

Phillips *et al* (1999) give the example of TQM and Benchmarking[37] which they claim are on the wane as contemporary business management methods, but are now being actively promoted in education. Another cogent example is Management By Objectives (MBO), which was *the* fad of the 1950s and 1960s (Spell 2000), being recycled as *Performance Management*.

Redefining Educational Management

In view of the caution towards generic management theory, and to an extent the apparent difficulty of applying it to an educational setting, a 'new' field of management theory is emerging; *Educational Management Studies* (EMS). A number of educational academics sought to clarify the position of educational management in a ESRC sponsored seminar series, *Redefining educational management* (Regard 2000), under the umbrella of the British Educational Management and Administrative Society (BEMAS). From this they have formed the *Standing Conference for Educational Leadership and Management* (SCRELM)[38], to promote Education Management Studies – EMS.

For these academics EMS is a distinct and irreducible academic field rather than a form of knowledge, or a sub-set of other educational research or business management studies (Fitz 1999). Much of the thinking behind the need for a specific field of EMS lies in the concern that educational management was becoming a narrow technical activity with short term objectives, which did not acknowledge or take in to account the broader values of the education system

---

[37] However the are some difference in the application of benchmarking between education and business
[38] See Ribbins (1999a) regarding the establishment and aims of SCRELM

(Glatter 1987, Bottery 1999, Fitz 1999,). For example, Bottery (1999), argues that educational management needs to go beyond what he perceives as generic management:

> Educational management describes an activity which includes but transcends the application of generic managerial or administrative techniques, for its practice has values and purposes which go beyond that of the business sector. If 'management' is interpreted in simple technical-rational terms, this larger picture is lost, and educators travel down roads towards destinations of which they may be unaware, or on which they do not desire to travel.
>
> (p. 301)

Many organisations which are 'generically managed' may well claim that their organisational aims transcend their management processes, and go beyond simply making a profit. For example, charities and other more philanthropic companies who contribute to their local population and the environment. Indeed those companies which support initiatives such as Education Action Zones and City Technology Colleges would also claims that their aims go beyond the 'bottom line'.

Bush (1999) points out that it was in the 1970's that the 'new' and specific educational management theories began to be developed. During this period there was much industrial conflict, which on the face of it was far removed from the values and principles being aspired to by the education system. It is therefore perhaps not surprising that an industrial or business style of management was unattractive to many in education. However, this was also a time when many of the 'modern' theories and management practices were being developed, which are very different to what would seem to be the assumptions made above by Bottery (1999). Much of the imperative for this approach was in response to both the climate of industrial conflict and the emerging 'Japanese effect'. Management academics, such as Peters[39], Kanter[40], Ouchi[41], were contributing to this new management culture, with its emphasis on values, such as quality and co-operation. In practice this led to the 'new management' of TQM and Learning Organisations, which is very much in tune with what is now being pursued by EMS.

There has always tended to be a degree of competition between the different fields of management, which may be to the detriment of the overall subject (Shafritz and Ott 1992). However, there is perhaps a danger of EMS becoming too insular and blinkered, and not learning from, or contributing to, the overall field of management. Ribbins (1999) outlines the aims of SCRELM which quite reasonably include identifying research priorities and commenting on research proposals, but also limiting its membership primarily to 'RAE returnable' staff who publish in educational management and leadership journals. To effectively preclude the vast wealth of

---

[39] Peters argued against the used of narrow rationalist management techniques.
[40] Kanter argued against a 'segmentalist' approach to management.
[41] Ouchi helped popularise the 'Japanese approach to management'.

experience, both academic and practitioner, from outside of this relatively small group of people, can only weaken the quality, breadth and strength of the arguments. In a similar vein Fitz (1999) makes a case for a specialised EMS journal:

> For a better understanding of the specialised interests - the body of knowledge which provides the distinctive intellectual quality of the field [EMS] – there is one way, I believe of identifying the discursive peculiarities and its boundaries produced, defended and invested in by field members and that is to look closely at the most prestigious journal in the field, *Educational Management and Administration (EMA)*.
>
> (p. 316)

Again there would seem to be dangers in this rather parochial approach which can be illustrated with reference to an example from the journal. Cutler and Waine (1999) in their paper, *Rewarding Better Teachers? Performance Related Pay in Schools* make no reference to the vast amount of research and evidence from the business sector. The issue of PRP has been has grappled with for many years, and indeed the evidence would strongly support their arguments.

Convergence - another school of management thought?

It would however be wrong to suggest that the EMS completely reject business management theory and approaches, indeed the terminology appears in much of the writings, as Fitz (1999) points out:

> "...the models of management which find their way into EMS, models such as Total Quality Management (TQM), Investors in People (IIP), 're-engineering' and 'leadership and followership', or, terms such as Human Resource Management (HRM) and 'competitive edge', to name but a few, have their origins in the field of production".
>
> (p. 314)

In terms of applying organisational theory, supporters of EMS such as Bush (1999), suggests that systems like Total Quality Management (TQM) and Human Resource Management (HRM) may have some value to education, but points out that they must be very carefully evaluated and only implemented, if appropriate. Southworth (2000) uses a 'learning organisation framework' to assess the development of primary schools, and significantly highlights the gap between the aspirations of the theory and the eventual practice. Systems such as Investors In People (IIP) and 're-engineering', are equally applicable to the service sector. Furthermore, Deming (1986), whose writings are at the heart of TQM points out that his management methods apply equally well to the fields of production and services (in which he includes education).

Some theorists appear to view the field of business management largely as a single homogeneous subject, based primarily on industrial production systems. However, it is in reality a rather disparate collection of theories, experiences and approaches. Shafritz and Ott (1992) point out that:

> There is no such thing as *the* theory of organizations. Rather, there are many theories that attempt to explain and predict how organizations and the people in them will behave in varying organizational structures, cultures, and circumstances
>
> (p. 4)

This therefore leads to the question: is educational management theory just another school of management theory? There would seem to be increasing evidence of cross-sector learning and the use and modification of theoretical models. For example, Bush and Coleman (2000) use a model[42] to describe school culture which looks very similar to the well know Boston Consulting Group's matrix (*Table 1*). This model was originally developed in the 1970's to analyse products in terms of market growth / share and to contribute to strategic planning processes (Kotler and Armstrong 1992);

Table 1: Product and School theory

BOSTON MATRIX (PRODUCTS)	BUSH AND COLEMAN (SCHOOLS)
Star	Moving School
Dog	The Stuck School
Cash Cow	The Promenading School
Problem Child	The Wandering School

In a similar way, Ouston (1999) compares the list of 'features of effective schools' from Barber *et al* (1995), with Peters and Waterman's (1982) work, and identifies a number of overlaps. Indeed she makes the point; "In fact Peters and Waterman's list might be more appropriate to schools in the current political context, in particular 'managing ambiguity and paradox', autonomy and entrepreneurship' and 'simultaneous loose-tight properties'." (p. 172).

Quite a number of educational theorists have used mainstream business theory to good effect, (eg. Morley and Rassool 2000, Tymms 1993, Cowie 2000), and as Hay McBer (2000) point out there is potentially much that can be learnt the 'other way'. In essence, there would seem little doubt that there is increasing convergence between the fields, and that will in the long term be to the benefit of both education and industry.

---

[42] Developed by Hopkins (1994)

Conclusion

This chapter sought to make the case for using theories which mainly originate from the business sector, (Organisational Theory and Management Accounting Theory), to help develop a theoretical basis for this thesis. In terms of business accounting it is argued that many of the problems and challenges which exist for businesses are very similar to those in educational accounting or performance measurement. Furthermore, businesses have been grappling with many of these issues, such as dealing with intangibles, for hundreds of years. Like the teaching of reading, the business world goes through fads and fashions, and frequently re-invents or re-cycles old ideas: and just as there is a danger of becoming permanently fixed to one model, there are also dangers to both sectors in continually jumping between models.

Overall, it is concluded that the key factors in educational management are essentially the same as those in business management; namely understanding and influencing human behaviour, whether this be in the case of teachers, accountants, nurses, or footballers. In both education and business, the term *Leadership* is very much in vogue, and as Drucker (1973) points out, organisations by definition are fundamentally human, not financial entities. From this perspective, Ranson and Stewart (1994) point out that: "Good managers have the same tasks and qualities whatever sector they are in". (P 26).

Indeed, similar observations had been made long before this. Shafritz and Ott (1999) point to Socrates, who in around 400BC, lists the skills of good managers and argues that they can lead any type of organisation, be it a choir, army, business or indeed family:

> Do not, therefore, Nicomachides....despise men skilful in managing a household; for the conduct of private affairs differs from that of public concerns only in magnitude; in other respects they are similar; but what is most to be observed, is, that neither of them are managed without men, and that private matters are not managed by one species of men, and public matters by another; for those who conduct public business make use of men not at all differing in nature from those whom the managers of private affairs employ; and those who know how to employ them conduct either public or private affairs judiciously, while those who do not know will err in the management of both.
>
> (Xenophon 1869 p.433, Cited in Shafritz and Ott 1999)

Chapter 4

The broader context and Organisational Theory

Politics is the art of looking for trouble, finding it, misdiagnosing it and then misapplying the wrong remedies (Groucho Marx)

Introduction

This chapter looks at the field of Organisational Theory (OT), to consider how the management context may have an influence on the impact of PIs. Chapter 2 emphasised that PIs do not exist in isolation, but are a part of, and interact with the broader management system. And chapter 3 suggested that the new field of EMS may in due course come to be recognised as a specific school of thought within OT. OT is a very large field which has developed over a long period of time, and this has led to many different schools of thought. Given the extensive use of OT in both this chapter, and throughout the thesis some clarification of the term is desirable. The term itself is relatively nebulous, and as well as overlapping with other fields, it is often used synonymously with other areas or labels such as; *Organisational Behaviour, Organisational Development, Management Studies, and Management Theory.*

With respect to the *Theory* part of OT, this has come to be used in the social sciences in what is technically a rather loose way. The 'theories' referred to are not theories in the proper scientific sense, such the 'the theory of relativity', but are based more on individual views, which may then go on to become generally accepted. This can be illustrated with reference to the *New Shorter Oxford Dictionary* (1997) which gives several definitions of theory including; (1) *a hypothesis that has been confirmed or established by observation or experiment and is accepted as accounting for known facts.* (2) *An unsubstantiated hypothesis; a speculative (esp. fanciful) view; an individual view, notion.* A great deal of organisational theory is of this second kind.

For Popper (1963) the key to distinguishing between science and pseudo-science was falsifiability (or refutability or testability). He pointed out that; "it is easy to obtain confirmations, or verifications, for nearly every theory – if we look for confirmations". (p. 36). And from this the key to the scientific status of a theory is that it is falsifiable, indeed he argued that a theory which is not so, is a 'vice'. Therefore, much of the organisational theory considered in this chapter can not be considered scientific theory in the proper sense, as it can be both vague and in some instances untestable.

With respect to the *Organisational* part perspective, Koontz *et al* (1980) discuss the problem of the ever widening field:

> Largely because of the influence of Chester Barnard and his broad concept of "organization" as referring to almost any kind of interpersonal relationships, it has become customary, particularly in academic circles, to use the term "organization theory" to refer to theory pertaining to almost any kind of interpersonal relationship. While many scholars attempted to make this field equal to management theory, it is now fairly well agreed that managing is a narrower activity and that management theory

> pertains only to theory related to managing. Management theory is often thought of as being a subset of organization theory, and it is now fairly well agreed that the general concept of organization theory is too broad.
>
> (p. 84)

In practice this has resulted in many areas of organisational or management studies being 'lumped' into OT, indeed much of this may well have little, if any, significant theoretical basis. A number of writers have picked up on this potential short coming. For example, Thompson and McHugh (2002) point out that OT is a massive field of study that is in danger of becoming a vague, and over-theoretical body of literature with little practical value. With this warning in mind, this and subsequent chapters, consider a limited number of the most widely used and enduring areas or aspects of OT. Furthermore, the empirical part of this thesis aims to contribute, all be it in a small way, to a more scientific approach to the use of management or organisational theories. And from this it is hoped that the research may contribute to the movement towards the adoption of evidence based policies in public management.

Most writers agree that there are many different approaches or schools of thought relevant to OT; for example, Shafritz and Ott (1992) identify eight groups of theories;

- Classical Organization Theory
- Neo-classical Organization Theory
- Human Resource Theory
- 'Modern' Structural Organization Theory
- Systems, Contingency, and Population Ecology Organization Theory
- Multiple Constituencies / Market Organization Theory
- Power and Politics Organization Theory
- Organizational Culture and Symbolic Management Organization Theory

(p. 8)

There is however little common agreement as how the many different theories and approaches should be grouped. Koontz (1961), in his book *The management theory jungle*, described management theory as a 'semantics jungle' and identified six groups of theories. In a subsequent book, *The Management Jungle Revisited*, (1980), he increased this to eleven. Most of the groupings, as with Shafritz and Ott's (above) are essentially chronological. This helps to show the development of management thinking, as the various theories tend to be cumulative, in that they build on previous work. However, it is also true that some theories have developed by rejecting other work, indeed as Shafritz and Ott (1992) point out there can be a degree of rivalry between the schools;

> Organization theorists from one school will quote and cite each other's works regularly. However, they usually ignore theorist and theories from other schools-or acknowledge them only negatively.

Other theorists have taken simpler views of the groupings. for example Pugh (1990) identifies two main schools, the *Organisers* and the *Behaviourists*. In a similar way, Wickens (1995) identifies, the *Controllers* and the *Behaviourists*. For the purposes of this thesis, the more complex groupings, such as the Shafritz and Ott model is considered to be too detailed, and the two dimensional models too simplistic; therefore a 'third way' or three group model is proposed. This comprises of three schools of thought which are classified as: *Rational / Scientific*, *Behaviourists*, and *Hybrid*.

To help put these in to context *Table 2* identifies the groups or schools, the theories considered in this thesis and some of key features:

Table 2: Summary of key Theorists and Theories

SCHOOL	THEORISTS / THEORIES CONSIDERED	KEY FEATURES
Rational / Scientific	Bureaucracy (Max Weber) Scientific management (FW Taylor) Planning (Henri Fayol)	Bureaucratic systems used to manage and control organisations. Assumption of a 'one best way'. Extensive use made of formal planning systems. PIs play a central part in the process.
Behaviourists	Hawthorne Studies (Elton Mayo) Theory X and Y (Douglas McGregor)	A reaction to the rational / scientific schools. The importance of the softer human aspects of system emphasised. PIs less important with issues of complexity and ambiguity being recognised.
Hybrid	Total Quality Management (TQM) Learning Organisations Business Process Reengineering (BPR)	Developed to meet the 'Japanese' effect, therefore PIs important. Combine aspects from the other two schools. May be 'hard' systems orientated, or 'soft' human orientated, or a combination of both.

Rational / Scientific school

The Rational /Scientific school which mainly originates from the early 20th century, represents the first systematic and detailed study of management and organisational behaviour. This is however not to say that management theory did not exist before. The church and military for example, developed many organisational features which persist to the present day, not only in these organisations, but in many others as well. Johnson and Scholes (1993) point out that the term 'strategy' originated from large military operations (the art of generals), which they trace back to Sun Tzu's classic treatise, *The Art of War*, written some 3000 years ago. More recently, Adam Smith's book *The Wealth of Nations* (1776), made an important study of the factory systems and division of labour. This has had a major impact on many private and public organisations up to the

present day. From a political perspective, Machiavelli's work[43], some two hundred years before this, which advocated the principle of 'unity of command'[44], and 'how to succeed at any cost', would seem relevant to the modern day political context.

This school (Rational / Scientific) is considered in terms of three key areas and early day 'management gurus'. Firstly, the work of Max Weber and *Bureaucracy*; then principles of *Scientific Management* from FW Taylor, and finally *Planning*, including the work of Henri Fayol. Although these theories developed independently there is a significant amount of overlap, indeed much of Taylor's and Fayol's work supports Weber's bureaucracy theory.

Bureaucracy - Max Weber

Max Weber (1864 – 1920) was a sociologist interested in the historical developments of civilisations through the studies of religion, economics and sociology (Pugh *et al* 1971). He was particularly interested in authority structures, and identified three types: **charismatic** – as the name suggests, someone akin to a modern day Richard Branson; **traditional** – where authority is inherited or passed on; and **rational / legal** – the basis of bureaucracy.

It is this third type, rational / legal, or bureaucratic approach which is of interest to this thesis. Although many of the observations made by Weber are aimed at the private sector, as Ostrom (1974) points they are also relevant to the public sector;

> Weber's theory of bureaucracy, was fully congruent with the traditional theory of public management in both form and method.
>
> (p. 9)

and as Hughes (1986) suggests more specifically for education;

> Schools and colleges, particularly if they are large, conform to a considerable degree to Weber's specification of bureaucracy.
>
> (p. 8)

The bureaucratic system was based on the detailed planning and organisation of the activities (*the rational*), and a collection of rules and procedures for those working in the organisation (*the legal*). In effect the bureaucracy aimed to ensured that the organisation was tightly controlled, and operated in the most efficient way possible. Weber formulated six principles for a 'modern' system of bureaucracy:

---

[43] In particular the Discourses (1513) and The Prince (1532). See also Harris (1999).
[44] A single person in charge is better than two more competent people together. This principle underlies much of the philosophy of present day Head teacher training (Grace 2000a).

1. The principle of fixed and official jurisdictional areas with detailed official duties. Conduct is limited by rules and which are methodically provided for each organisational activity. The authority is with the post, rather than the individual.
2. A firmly ordered hierarchy of superior and subordinate offices.
3. The management of the modern office is based upon written documents ('the files'), which are preserved in their original or draft form.
4. A prescribed course of expert training which prepares the individual to hold office.
5. The holding of the office is the primary activity of the office-holder who is rewarded by salary. (This aims to replace a system of appointments based on patronage or nepotism)
6. The conduct of each office, and relations between them, are covered by the general rules.

(adapted from Jones and Dugdale 1995: 301)

From a present day perspective, bureaucratic organisations are assumed to have the following features: Staff are permanently employed (ie contracted), they have undergone appropriate professional training and are a part of a hierarchy with clear lines of responsibility. Their relationship with the organisation and other employees is defined by contract (both implicit and explicit) and their role and specific tasks in the organisation are clearly defined.

Originally the word bureaucracy was used in a positive way to describe a well run and efficient organisation (Berry *et al* 1995), indeed it was a matter of pride to be called a bureaucrat. However, nowadays it is much more a term of derision, and used to describe an inefficient organisation, which is cumbersome and inflexible. In short, bureaucracy has come to be used to describe what is wrong with the 'old' management, and what the 'new' management aims to replace. Drucker (1976) highlights this popular perception;

> Public service institutions are prone to the deadly disease of 'bureaucracy'; that is towards mistaking rules, regulations, and the smooth functioning of the machinery for accomplishment, and the self-interest of the agency for public service.
>
> (p. 12)

More specifically in education Bush (1989) points to the tension between a bureaucratic system wanting to impose control, and the creativity of the professionals working in the system:

> The pervasive influence of bureaucratic and other rational models on schools and colleges is confirmed by the fact that all the other perspectives tend to be tested against the bureaucratic 'norm'. A major criticism of this approach is that it neglects the individual qualities of people and regards them as part of the organizational structure, slotting into defined positions in the hierarchy. Schools and colleges are staffed mainly by professionals who require substantial discretion in performing their teaching role. The bureaucratic model does not satisfactorily explain the contribution of professional staff to the management of educational institutions.
>
> (p. 5)

Furthermore, in terms of PIs bureaucracies can have an undesirable depersonalising effect, with the responsibility for an individual's performance lying with the 'post' rather then the particular

person. Accountability systems, such as individual appraisal, attempt to overcome this effect. Hughes (1994) claims that whilst there has been a move away from bureaucratic management in both the private and public sectors, the main difference is that bureaucracy has lasted a lot longer in the public sector;

> The earlier, administrative, rigidly bureaucratic model is now discredited both theoretically and practically. A new model of public management using theories derived from economic and private management has taken over and will totally change the way the public sector operate. Public management in the next century will not be the rigid, bureaucratic kind in place for most of the twentieth century. It will pay attention to results above everything else, to flexibility rather than rigidity and to the political context in which the public sector operates....but there will be no going back to the traditional model.
>
> (p. 22)

It is often claimed by politicians that many of the reforms in the public sector are aimed at reducing bureaucracy, indeed such claims receive almost universal approval. There are many examples in education of bureaucracy; for example, the obvious hierarchy, staff employed on contracts, expert training and clearly defined methods of working. However, this is not necessarily bad. There can be benefits to those both working in, and those receiving the services; for example, few would support the notion of a school without a permanently contracted trained head at the top of the hierarchy. The same is true in other areas, for example, it is comforting to know that the air traffic control systems is organised along bureaucratic lines. Jaques (1991) in his paper[45], *In Praise of Hierarchy*, considers the rather unfashionable view that there might be some benefit from bureaucracy and hierarchies;

> At first glance, hierarchy may seem difficult to praise. Bureaucracy is a dirty word even among bureaucrats, and in business there is a widespread view that managerial hierarchy kills initiative, crushes creativity, and has therefore seen its day. Yet 35 years of research have convinced me that managerial hierarchy is the most efficient, the hardiest, and in fact the most natural structure ever devised for large organizations. Properly structured, hierarchy can release energy and creativity, rationalize productivity, and actually improve morale. Moreover, I think most managers know this intuitively and have only lacked a workable structure and a decent intellectual justification for what they have always known could work and work well.
>
> (p. 255)

The key point here is that there is good and bad bureaucracy. In a similar vein Mintzberg (1993) distinguishes between *Machine Bureaucracy* and *Professional Bureaucracy*. The former involves a low degree of strategic involvement by individuals, whereas the latter allows far greater involvement. Despite various claims and political aspirations, there is evidence of increasing bureaucracy (in particular machine bureaucracy) in education. A number of the 'modern' developments in management, in particular those under the guise of accountability, are very bureaucratic: for example, the Ofsted inspection process, performance management and the GTC's

---

[45] Originally in *Harvard Business Review* (Jan-Feb 1990)

52 rules of conduct (Mansell 2000). Research by Cowie (2000) on the relationships between heads in Aberdeen and the EA, points to a mismatch in expectations and actuality, with the heads generally wanting to be more involved in the strategy formulation.

It is also questionable whether other sectors, including the private, are rejecting and moving away from bureaucratic structures. Although many cases can be cited of organisations adopting less bureaucratic structures, and this being of benefit to both the organisation, and its customers, the success to a large degree depends on the particular industry. Indeed, for many situations, bureaucratic approaches can be of benefit. In medicine, for example, it can be argued that bureaucracy saves lives; for instance, expert training, combined with set protocols, allow non specialists (eg paramedics) to significantly improves survival rates following heart attacks.

Even some of the new DotCom industries are adopting bureaucratic principles. For example, in the case of on-line banking the traditional hierarchical structure with the manager at the top, has been replaced by a 'virtual' hierarchy, which is governed by structures and procedures. Methods of working are precisely defined or scripted, and anything that departs from the norm is 'passed on'. Indeed, one wonders what Weber might have written, had he spent hours on the phone, going round in circles trying to talk to a human (with a name) in a call centre, whilst listening to the four seasons. As Jaques (1991) points out much depends on the particular implementation of bureaucracy;

> As presently practiced, hierarchy undeniably has its drawbacks. One of business's great contemporary problems is how to release and sustain among the people who work in corporate hierarchies the thrust, initiative, and adaptability of the entrepreneur. This problem is so great that it has become fashionable to call for a new kind of organization to put in place of managerial hierarchy, an organization that will better meet the requirements of what is variously called the Information Age, the Services Age, or the Post Industrial Age.
>
> (p. 225)

Whatever the claims about new methods of working, it would seem that bureaucracy will continue to form much of the bedrock for organisational structures in education, and many other organisations. However, this is not to say that the 'new' managerial forms will not also play an important part; far from it, however, it would seem that this will still be with in a bureaucratic framework for the foreseeable future. And as Jaques (1991) points out the key is understanding these structures:

> Hierarchy is not to blame for our problems. Encouraged by gimmicks and fads masquerading as insights, we have burdened our managerial systems with a makeshift scaffolding of inept structures and attitudes. What we need is not simply a new, flatter organization but an understanding of how managerial hierarchy functions - how it relates to the complexity of work and how we can use it to achieve a more effective deployment of talent and energy.
>
> (p. 255)

Scientific Management - FW Taylor

Fredrick Winslow Taylor (1856 – 1917) is synonymous with the principle of scientific management, which aims to find the *one best way* of doing a particular task. Taylor joined the Midvale Steel Works as a labourer without any qualifications, and through night school gained a degree in engineering, and progressed to become Chief Engineer. He saw many aspects of the organisation, and this had a significant effect on his views on workers and management. Taylor is commonly referred to as the 'father of scientific management', indeed the terms Taylorism and Scientific management are used interchangeably. In 1911 he produced his seminal work, *The Principles of Scientific Management*, which has had a vast impact on all sectors of industry in many countries as Huczynski and Buchanan (1991) point out:

> Taylor's writings continue to have an enormous influence on management practice. In many organizations today, work is organised in the ways that Taylor suggested.
>
> (p. 281)

Perrow (1979) too emphasises the significant influence on the organisation of work that scientific management continues to have;

> These principles (scientific management) have worked and are still working, for they addressed themselves to the very problems of management, problems more pressing than those advanced by social science.
>
> (p. 79)

Indeed, as Merson (2000) points out many of the 'modern' management techniques, (eg. performance related pay systems), have been influenced by Taylor's work;

> Behind the rhetoric of modern management lies the architecture of hierarchical management adopting processes that are redolent of very traditional, authoritarian and Taylorist regimes.
>
> (p. 165)

Although scientific management is closely associated with factories and production lines, there is much evidence of his principles being applied to the service and public sectors (Hartley 1990, and Helsby and Saunders 1993). In essence Scientific management aims to find the best way of doing a particular task, by using time and motion studies, and by selecting the most suitable worker for the job. Previously work had been largely organised on a 'craft control' system, where expert craftsmen used their experience and a principle of rule of thumb to organise work. Taylor (1947) argued that the whole country (US) was suffering through inefficiency in almost every sphere. The solution to this was through and detailed systematic approach to management, in effect making

management a true science. To achieve this Taylor defined four underlying principles of management:

1. *The development of a true science of work*: This removed much of the uncertainty and vagaries of the design of the work, thereby allowing the workers to understand what was precisely required of them. It established the principle of defining what a suitable worker should achieve under optimum conditions, also connected to this was the notion of linking performance and pay.

2. *The scientific selection and progressive development of the workman*: This placed the responsibility on the management for selecting the most suitable men, in terms of physical and mental properties, for the particular job. Once selected, again the management were responsible for providing systematic training, which would enable them to achieve the best possible performance. It was Taylor's belief that if the right men were selected they could all be trained to perform well.

3. *The bringing together of the science of work and the scientifically selected and trained men*: In this Taylor argues that much of the resistance to scientific management comes from the management, with workers in general being willing to learn and to do a good job, which would lead to higher pay. Taylor refers to the need for a 'mental revolution' on the part of management to accept and adopt the new principles.

4. *The constant and intimate co-operation of management and men*: This established the principle of the division of work, with managers specifying the methods, carrying out quality control and supervising the workers. Significantly, Taylor emphasised the need for close co-operation, which would reduce the opportunity for conflicts; furthermore, managers themselves would be under the scrutiny from the workers.

(Based on Taylor 1947)

In addition, with respect to organising factories, Taylor introduced the concept of functional management. This divided up the many different functions of the foremen which were then carried out by specialists, such as, the cost clerk, time clerk, repair boss, inspector and shop disciplinarian (Pugh *et al* 1971).

In education and the public sector, there are as Helsby and Saunders (1993) point out, many examples of the adoption of scientific management principles;

> It is easy to identify elements of Taylorism in the operation of many of their large-scale social and educational programs. Thus there is a recurrent pattern of centrally devised and standardised programmes, divided into manageable units, aimed at particular target groups, implemented by teacher/workers, who had not been involved in the planning of the programme, constantly monitored by outsiders concerned with efficiency, and finally judged by 'consumers' rather than by workers.
>
> (p. 56)

In practice, the National literacy strategy (NLS) provides a good example of a centrally devised and standardised programme, with the requirement[46] to follow the set method. Teachers are generally selected along relatively 'scientific principles', and are in theory at least, rewarded in terms of their

[46] Although not a legal requirement.

performance. The principal of functional management too can be found in most schools, with Heads of Years and Departments, and PIs and league tables providing the means by which the consumers are encouraged to judge success.

Furthermore, Hartley (1990) suggests that much of education policy is based on pseudoscientific management principles, to create an 'air' of precision and neutrality, but may in fact be used to hide political and ideological dogma;

> ... the terminology used in recent education policy documentation contains a scarcely-disguised appeal to the tenets of Taylorism and its metaphorical basis, namely the machine...the power of metaphors in structuring education policy and practice is considerable...in a rational and technical society they appear to give scientific status to the contexts in which they are used. Their facade of certainty and political neutrality gives a spurious objectivity to what are, at root, political and ideological means and ends.
>
> (p. 69)

The key to this is in the selection or rejection of the evidence, and as Hartley points out this is to a large degree a political issue. Scientific management has over the years been subject to intense criticism, and provides for many business management courses, the example of how *not* to manage. The Economist (1993) illustrates this common perception[47];

> Taylorism is now vilified as the epitome of a hierarchical, authoritarian style of management which caused decades of labour strife. No right-thinking manager today would describe himself as a disciple of Taylor.
>
> (p. 77)

Scientific management, particularly with its connotations of factories is often perceived as reducing the role of workers to that of 'cogs' in the systems. O'Neil (1986) suggest, that engineers like Taylor who became managers, tended to view workers as resources alongside raw materials and machines. There is evidence to support this view, with for example Taylor emphasising the selection of workers based on essentially their physical attributes.

Kelly (1982) points to the received wisdom that the main problems with scientific management is that it largely ignores the personal differences between workers, decreases the skill required for a particular job, with a consequential loss of status, and puts too much pressure on individuals to perform. There was substantial hostility from workers towards scientific management, so much so, that the US congress banned Taylor's methods from being used in the defence industry at the time of war[48] (Aitkin 1960).

---

[47] This is a selective quote from an article which looked at new approaches to scientific management.
[48] 1914

A vast amount of folklore has built up over the years with regard to scientific management, some of which is probably unjustified. Much of the problem would seem to lie in the implementation of 'scientific methods', combined with unreasonable and unrealistic claims about the benefits; in effect not thinking through and managing the change carefully enough. Indeed, in his later writings[49] Taylor (1916) highlighted this issue:

> I must speak of the fakers, those who have said they can introduce scientific management into a business in six months or a year. That is pure nonsense. There have been many strikes stirred up by that type of man.
>
> (p. 72)

Although there is perhaps a common perception that scientific management methods drive a wedge between manager and workers by emphasizing the different roles, Taylor did emphasise the need for both parties to work closely together, and that this would lead to mutual respect and feedback. A number of 'modern' British organisations, might not measure up well with these principles. Furthermore, central to Taylor's approach was the belief that scientific management would be, and should be, to both the benefit of the company and its workers, as well as society being rewarded with cheaper goods and services. Many (for example, Monbiot 2000) would argue that a significant proportion of companies nowadays would not take such a 'generous' view.

In terms of managing and controlling large and dispersed organisations such as the education system, a scientific management approach with the associated PIs can be of benefit. PIs can be developed to measure common processes across the system, from which valid judgements can be made, (assuming other important factors are controlled for). However, the crucial issue is how these and indicators are used and for what purpose. For example, most people would agree that performance information should be used to improve practice, ie. be the basis of *evidence based policies*. Indeed, it would seem that such an approach to practice and policy would have appealed to Taylor (1916):

> Scientific management makes no pretense that there is any finality in it. We merely say that the collective work of thirty or forty men in this trade through eight or ten years has gathered together a large amount of data. Every man in the establishment must start that way, must start our way [ie. as an apprentice], then if he can show us any better way, I do not care what it is, we will make an experiment to see if it is better. It will be named after him, and he will get a prize for having improved on one of our standards. This is the way we make progress under scientific management. There is your justification for all this. It does not dwarf initiative, it makes true initiative. Most of our progress comes through our workmen, but comes in a legitimate way.
>
> (p. 80)

---

[49] Address given on the 3 March, two week prior to his death

However, in present day policy formulation in education and the broader public sector, evidence is only one, and often it would seem minor factor in the process. Indeed, rather than being evidence based, the process is increasingly referred to as *Evidence informed* or just *Evidence aware* (Davies *et al* 1999). Leicester (1999) identified a number of factors which inhibit the adoption of evidence based practice, including politics and the effect of the bottom line (ie. KPIs). In practice this can lead to frustration on the part of practitioners, who may be able to demonstrate with the use of indicators the value of a particular method, but this may not be adopted because of other (eg. political or ideological) factors.

In both the private and public sector, Taylor's contribution to management has been substantial. Whilst there are problems and shortcomings with the underlying principles of scientific management, much of this comes down to selective and poor implementation, which would often appear to be the result of a narrow view of the fundamental aims and purposes of the organisation. From Taylor's groundwork others have developed and improved 'scientific management'; for example, Frank Gilbreth and Henry Gantt.

Planning - H Fayol

This section looks initially at the work of Henri Fayol, and then more generally at the area of planning. Fayol (1841 – 1925) like Taylor was by training an engineer, who also moved from a technical position into a managerial role. He made a number of significant contributions to the techno / rational management model, including the principle of hierarchies and rational planning. Fayol produced a seminal work, *Administration Idustrielle et Générale - Prévoyance, Organization, Commandement, Coordination, Contrôle* (1916)[50], which has had a major influence on many organisations.

Berry *et al* (1995) point out that Fayol separated the role of the managers from the managed (division of labour), and proposed that organisations should have a unified set of purposes, with individual interests being subordinate to these. Fayol's model of control ensured that communications followed a strict hierarchical 'up and down' structure, although sideways communication was allowed, so long as the appropriate line manager knew. Fayol proposed that workers should be responsible to just one manager ('unity of command'), which contrasts with Taylor's principle of functional management (Huczynski and Buchanan 1991). However, given Fayol's principle of sideways communications the differences may not have been so great, and much would depend on the particular situation or industry. Fayol defined management as comprising of five elements:

---

[50] Translated in to English 1949, with the title *General and Industrial management*

1. To forecast and plan: 'examining the future and drawing up the plan of action'.
2. To organize: 'building up the structure, material and human, of the undertaking'.
3. To command: 'maintain activity among the personnel'.
4. To co-ordinate: 'binding together, unifying and harmonizing all activity and effort'.
5. To control: 'seeing that everything occurs in conformity with established rule and expressed command'.

(adapted from Pugh *et al* 1971)

In terms of this thesis Fayol's most significant contribution, was the bringing together and formalising the principles of the rational planning systems. Clearly planning as such had occurred before this time, but it was Fayol's work which integrated these into the management process. Pugh *et al* 1971 point out that from Fayol's point of view, an effective organisation needs to have a plan which provided; *unity, continuity, flexibility and precision*. In essence, Fayol believed that the future could and should be planned for in detail, but whilst still maintaining a degree of flexibility:

> For Fayol, managing means looking ahead, which makes the process of forecasting and planning a central business activity. Management must assess the future and make provision for it.
>
> (Pugh *et al* 1971: 61)

Over the years many theorists and practitioners have contributed to the very broad field of planning. The remaining part of this section looks at some of these writings, including those specifically aimed at education. Firstly, the work of Koontz *et al* (1980) who identified four main uses of planning is considered;

1. *To Offset Uncertainty and Change*:- Organisation can not 'navigate' by aiming directly for goals alone, plans can provide useful 'way-points', to indicate the progress towards the goals.

2. *To Focus Attention on Objectives*:- The value and importance of interdepartmental contacts when producing the corporate plans.

3. *To Gain Economical Operation*:- Reducing piecemeal and uncoordinated activities, which will result in more efficient use of resources.

4. *To Facilitate Control*:- Setting and defining standards upon which the subordinates' accomplishments can be measured.

(p. 170)

Although this was written with commercial businesses in mind, it is equally relevant to education. Within schools PIs are used extensively to set 'way-points' or intermediate targets. Individuals, departments and schools bring together objectives (in the form of PIs) to produce 'bigger' plans, and these help coordinate the various activities. Finally, the plans and the associated PIs are used extensively as control mechanisms (Broadbent *et al* 1999, Cowie 2000).

Sizer (1982) differentiates between planning (plain and simple) and strategic planning. The former deals with the here and now, and essentially allocates resources; whereas strategic planning is more outward looking, extends beyond one year, and is able to set the direction of the organisation. It is

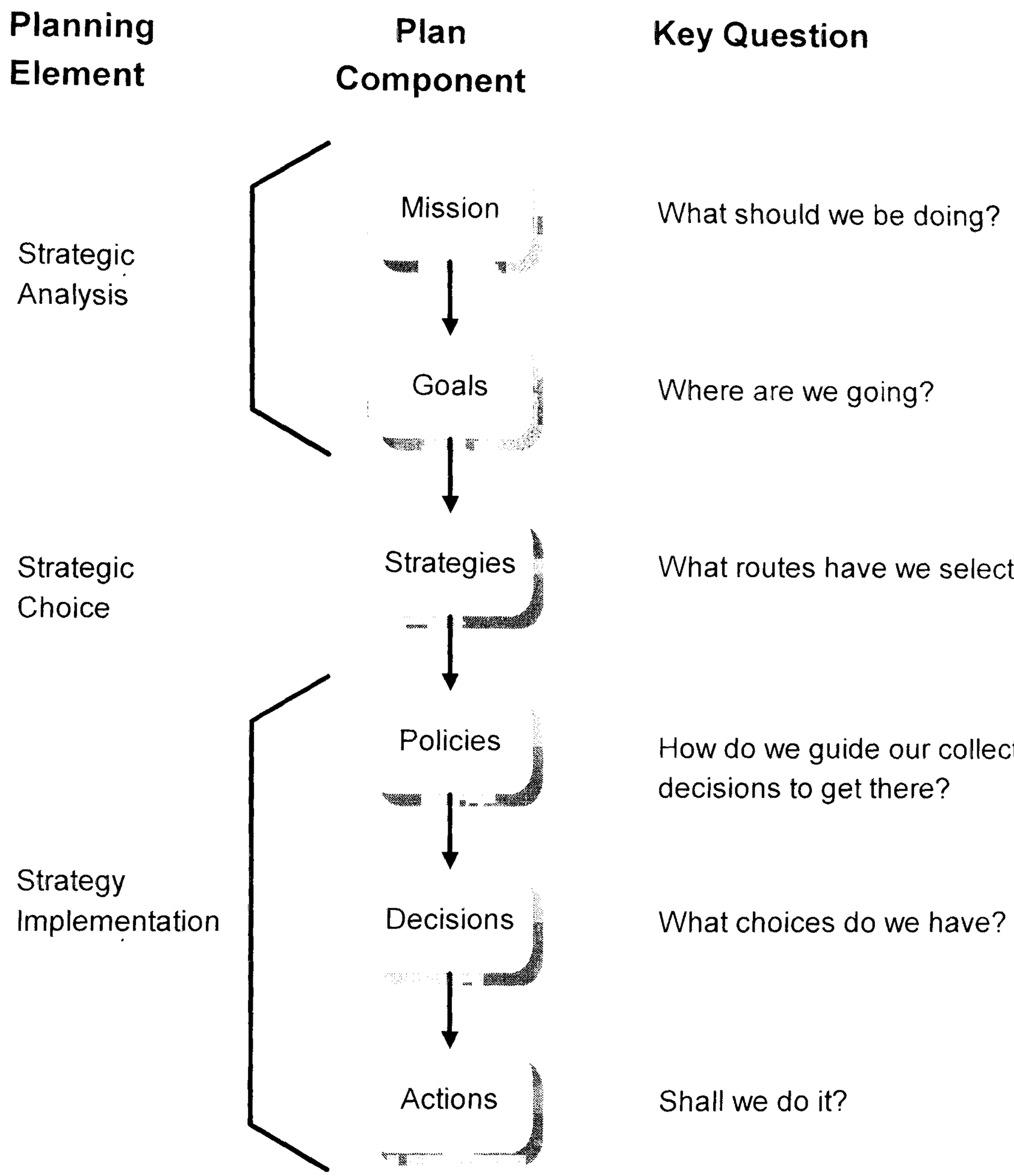
however, only over the last thirty or forty years that strategic planning, as a specific activity, has become widely used in business. Ansoff's (1965) book *Corporate Strategy* marked the beginning of a rapid growth in the study and application of strategy, although this is not to say that previously organisations did not carry out long term planning, or act strategically: as stated earlier the term strategy can be traced back thousands of years.

In a similar way, Fidler (1996) differentiates between school development planning and strategic planning, although as argued earlier in this chapter there is some doubt as the opportunity for meaningful strategic planning at the school level. In a similar vein, Bell (1998) suggests that: "the term strategic has become almost inseparable from the word planning such that if planning is not 'strategic' then it is not regarded as planning in any real sense." (p. 449).

In practice the different planning activities in schools would seem to fall somewhere between the two extremes, (short-term resource allocation and long term strategic planning). However, Bennett *et al* (2000), when looking at planning in effective primary schools found; "that 'planning' and 'strategy' are not conterminous for schools" (p. 336). This rather surprising finding they attribute to the effect of the Ofsted inspection process. In other organisational settings many theorists have highlighted the dangers of a disjointed approach to: strategy, planning and action (eg. Rangone 1997, Anthony 1988, Johnson and Kaplan 1991).

Most planning systems follow a broadly similar pattern: starting with the aims of the organisation, and leading logically to the choices and actions which will achieve those aims. *Figure 3* shows a typical 'generic' planning model;

Figure 3: A generic strategic planning model



Generic strategic planning model (Robson 1994)

In education, planning of some sort has of course always existed, however until the reforms of the 1980s this was primarily 'resource allocation'. Schools operated in a relatively predictable environment; the number of children joining the school was known, and the area of greatest cost, the teachers, was the responsibility of the LEA. Edwards (1996) points out that before 1993 school development plans were simply 12 month forecasts written by the head and 'nodded' through by the governors. However, as a result of the 1988 Act, planning became a far more important and critical part of the management process. The curriculum and budgetary reforms demanded new and

more sophisticated planning systems, which she describes as being initially 'reactive', in that they were designed to meet the 'letter of the law', rather than being genuinely strategic. However, once the reforms began to bed down, then planning could become more strategic and outward looking. The DES (1991) document, *Development Planning*, illustrates the broad requirements of the 1988 act:

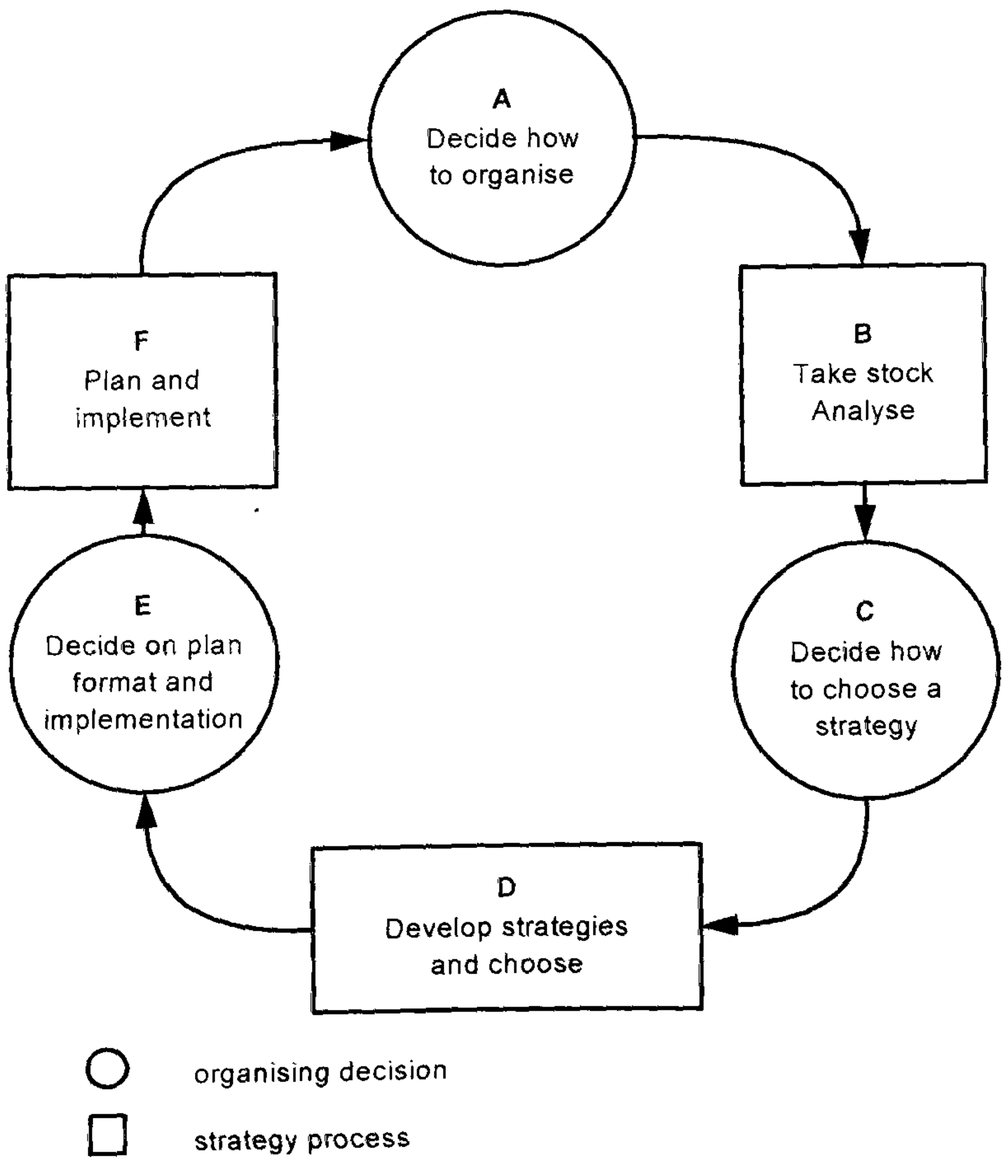
> A school development plan is a plan of needs for development set in the context of the school's aims and values, its existing achievements and the national and LEA policies and initiatives. Detailed objectives are set for one year: the objectives for later years are sketched in outline. The purpose of development planning is to assist the school to introduce changes successfully, so that the quality of teaching and standards of learning are improved.
>
> (p. 2)

From this Hargreaves and Hopkins (1991) identified four main stages in the planning process:

> - *Audit:* a school reviews its strengths and weaknesses
> - *Construction:* priorities for development are selected and then turned in to specific targets
> - *Implementation:* the planned priorities and targets are implemented
> - *Evaluation:* the success of implementation is checked
>
> (p. 4)

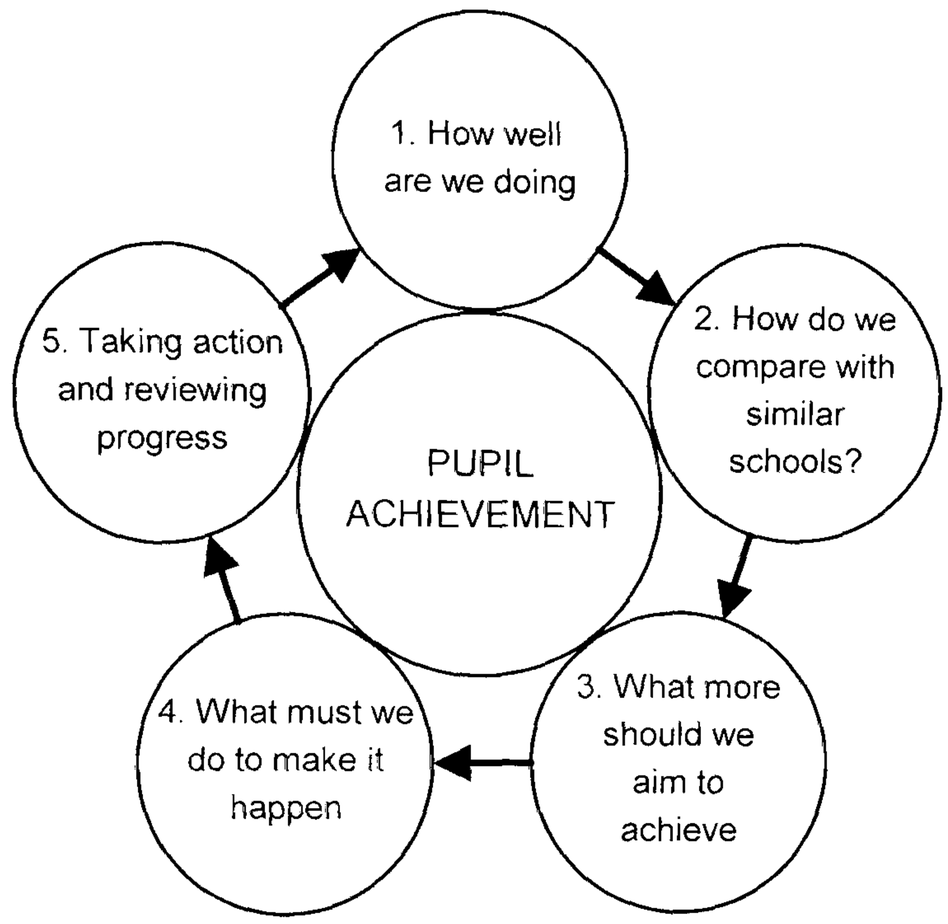
The model below (*Figure* 4) from Fidler (1996: 57) essentially follows these principles and is very similar to the previous generic model (*Figure 3*)from Robson:

Figure 4: A basic model of school strategic management



In practice many examples based on this generic model can be found in the education system. The model below is from the *Autumn Package* (DfEE 2000b) and is used to plan school improvement:

Figure 5: Five Stage Cycle for School improvement



Performance indicators are a significant part of planning systems. In the example above PIs will be used extensively, for instance, *Benchmarking* may be used to compare schools. However, if this process is flawed, (eg. the comparisons are not valid), this may well have a negative effect on the rest of the planning process, for example, this might encourage the school to set inappropriate targets (at stage 3).

Although there would probably be general agreement that planning is of value, currently the education system abounds with plans at all levels, ranging from individual (pupils) education plans, to DfES strategy documents. This is in spite of claims of reducing administration and bureaucracy. Indeed, the government's view appears to be that there can never be enough planning; furthermore, all parts of the system are largely judged on their ability to produce 'impressive' plans, whether or not they have any real meaning, or are of any use to those writing the plans or receiving the service. Davies and Ellison (1998) succinctly summarize the phenomena: "The thicker the plan the less it affects classroom practice". (p. 461), and this may encourage 'lip service' towards plans, and suspicion of the PIs contained in them.

Whilst the education system (and much of the rest of the public sector) have been increasing the number and more significantly the apparent importance of planning, there has been something of a counter movement in some parts of the 'business world'. Well before the explosion in educational planning, Lawrence and Lorsch (1967), argued that formal planning tended to restrict an organisation's competitiveness in a growing and changing environment, largely by discouraging flexibility. In a similar way, Mintzberg (1994) makes the point that strategic planning by its very nature, is prone to developing into an overly rigid focus on analysis and quantification; making it innately inflexible and incapable of predicting crucial market shifts or encouraging appropriate adaptation by the organisation. Perrott (1995) pointed to the problems of planning having to follow arbitrarily set fixed cycles, in particular financial reporting periods, (Ofsted inspection cycles would also seem relevant). Hamel and Prahalad (1994) argue that rational planning models fail to allow or encourage the creative processes and discoveries that generate breakthroughs and prompt genuine shifts in strategic direction. (Their work which is at the heart of BPR is considered later in this chapter).

On the other hand there are also dangers in a very flexible or loose planning process. Wallace and McMahon (1994) point out, that whilst 'day-to-day crisis management' is very flexible, it is likely to reduce efficiency and effectiveness. Hughes (1994) discusses many of the criticisms against strategic planning and concludes that on balance, used carefully, they can be of benefit. The potential benefits include helping facilitate the 'thinking processes', and encouraging and enabling the involvement of all staff in the management of the organisation (see also Wallace and McMahon (1994).

Bell (1998) argues for a more flexible and synergistic approach to planning and management, and one which encourages organisational learning[51]. Glatter (1999) points to the contradiction, with on the one hand ever increasing techno-rationalist systems such as strategic planning, and on the other, an increasing awareness of the turbulence and complexity of the educational environment. In practice, Davies and Ellison (1998) point to the difficulties with formal planning systems in situations of complexity and ambiguity:

> Strategic planning may well be useful for the more predictable and controllable elements within the planning processes, especially when these are incremental and linear and where a good understanding of the detail is possible. We contend, however, that there are problems associated with this approach. It assumes a rationality and predictability which, in practical terms, may not be possible in a turbulent, dynamic environment such as we have now. Its incremental nature can mean that, instead of taking a fresh zero-based view of the planning needs of the next five years, the present plan and its associated culture lead to a focus on adjustments and minor developments. Thus, strategic planning and development can obscure or overtake the need for more radical strategic rethinking.

---

[51] In line with the 'learning organisation' concept which is discussed later in this chapter

(p. 462)

Whilst highlighting the limitations of the current formal systems for schools, they also take the issues beyond the problems of rational planning and into the realms of 'futures thinking' (see Davies and Ellison 1999) and more radical approaches to management, such as BPR.

Planning is clearly an important process in the management of complex systems, such as education; however, many of the problems with rational planning systems and PIs centre on the fact that concrete predictions in terms of outcomes have to be made. If the planning processes themselves are flawed or found wanting, so to will be the predictions or targets. And to hold people accountable for these may well have negative effects on both their performance and that of the system as a whole.

Behaviourist School

The rational / scientific approach to work and organisations improved the measurable efficiency of many organisations; in education too, scientific methods, such as the National Literacy and Numeracy Strategies (NLNS) has raised apparent standards[52]. However, the most fundamental criticism or weakness is the lack of regard for individuals in the system, which may in turn affect the effectiveness of organisations, particularly complex ones which rely on human interactions. The Behaviourist (or human relations) school of management aimed to address this short-coming. Perrow (1979) succinctly points to these two different approaches:

> Whilst Taylor [*Scientific Management*] had been aware of group solidarity he seems to have seen it as an obstacle to be overcome rather than as phenomenon which needed to be understood and turned to management's advantage. For human relation theorists, however what was required for a smoothly functioning organisation was no less than a rational assessment of the whole person, set in a context of the social relations of the workplace.
>
> (p. 64)

The work of Henry Gantt, who is credited with the 'humanisation of Taylorism' (Huczynski and Buchanan 1991) provides a useful link between the two schools. He worked with Taylor at the Bethlehem Steel Works[53], and whilst endorsing the principles of Scientific management, he was concerned that the principles were used oppressively and as an instrument of control:

> Gantt realised that the worker was a human being with needs and dignity that deserved consideration by management....Nevertheless (like Taylor), he believed that the opportunity to earn money was all the motivation the workers need to accept the improved methods. Management's job was to create the conditions in which this could happen.

---

[52] Although there is an increasing body of evidence which questions this rise (Tymms and Fitz-Gibbon 2001)
[53] It has just been announced (November 2002) that this is about to close

(Huczynski and Buchanan 1991: 289)

Gantt's system replaced the term 'one best way' with 'best known way at present', and like Fayol proposed that line management should not be organised on a functional basis, but rather that supervisors should be responsible to for all aspects of the supervision and management of 'their' workers. Gantt also proposed that bonus system should recognise all of the contributors, and not just one individual. In practice this means rewarding other team members as well as the supervisor. This approach is very common in industry today, although not generally so in education. Teachers individually may pass the threshold and receive a 'bonus', however the results upon which they have been judged, may in part be due to the efforts of others. It is impossible to precisely allocate the contribution of individuals in a complex organisation. In spite of these 'improvements' the behaviourists took a fundamentally different view of organisations, namely from the perspective of those working in them.

Elton Mayo

Elton Mayo (1880-1949) has been referred to as the founder of both the Human Relations movement and of industrial sociology (Pugh 1971). However, it is his contribution to the Hawthorne studies which started in 1927, for which he is best known. As Shafritz and Ott (1992) point out this research has, even up to the present day[54], had a major impact on management thinking:

> The Hawthorne studies laid the foundations for a set of assumptions that would be fully articulated and would displace the assumptions of classical Organization theory[55] twenty years latter.
>
> (p. 144)

Mayo and his team studied the effect of different working conditions on the output of the workers at the Hawthorne factory. Initially, they looked at the effect of different levels of light and surprisingly it was found that both increasing *and* decreasing the levels led to an increase in production. What followed was five years of research in to many other factors which it was felt could affect production. These included, varying the length of the working day, the time and length of breaks, giving individual and group incentives; and virtually whatever was changed led to an increase in production by the experimental group. However, as Parsons (1978) points out, a certain degree of mythology has sprung up around Hawthorne, and therefore some of the findings should not be taken too literally. For example, he suggests that the effect of workers improving their skills over time was not properly controlled for, and this would have contributed to the improved outputs.

---

[54] His research is still subject to scrutiny and re-evaluation (eg. Parsons 1974 and 1978, Adair 1984 and Diaper 1990)
[55] Essentially what is referred to in this thesis as Rational / Scientific

Mayo's observed that 'his' experimental group experienced an increase in their work satisfaction, which he attributed to the high level of social interaction amongst the group themselves and with the researchers. This he concluded led to the substantial increase in productivity (Roethlisberger and Dickson 1964). The effect of increasing performance because of particular interest being shown in individuals and groups has become known as the *Hawthorne Effect*.

Examples of the Hawthorne effect can be observed in many areas of education, for instance attendance and truancy rates. A school or teacher taking an interest in this, and measuring performance, will not surprisingly lead to an improvement. Conversely, choosing to ignore them may well have the opposite effect. In a similar way it would be expected that concentrating on a specific group of pupils, for example with the 'booster class' system[56], would lead to an improvement in their KPI performance. At the local area level, initiatives such as 'Beacon schools' would probably have a similar effect, and help raise that schools KPI performance.

On the other hand schools which are struggling and nearly failing may experience an 'anti – Hawthorne effect', with very little interest will be shown in them. The LEA and the school itself may well want to keep a low profile. However, if they go on to be declared failing, suddenly great interest is shown. The staff will be encouraged to 'save' the school, and any glimmer of hope will be seized on and nurtured. This may help explain the seemingly miraculous and rapid improvements of some schools in such situations. It is suggested that this is not only simply because of; a new plan, a taskforce, or a new head, but also or perhaps more importantly, because of increased contact and feedback to the school and those working in it.[57]

The importance of these softer aspects are at the heart of Mayo's findings. As Koontz *et al* (1980) point out, it was not so much that these ideas were new, but rather that the Hawthorne studies demonstrated and provided compelling evidence to support the value of such approaches:

> What the Hawthorne studies dramatized was that humans are social - that business operations are a matter not merely of machinery and methods but also of gearing these with the social system to develop a complete sociotechnical system. These experiments led to increased emphasis on the behavioural sciences as applied to management and to the recognition that managers operate in a social system. It should not be inferred from this that prior to the Hawthorne experiments successful managers did not recognize the importance of the human factor, or that management theorists overlooked it... what the work of Mayo and his associates did underscore was the need for a greater and deeper understanding of the social and behavioural aspects of management.
>
> (p. 51)

---

[56] Pupils close to level 4 or 5+ A-C's

[57] The accompanying increased resources does of course also help

To help put Mayo's findings into a work or organisational context. Lawton and Rose (1991) provide a useful summary of the main points:

- Informal group dynamics are important
- Informal group norms can develop which are able to undermine formal working relations
- Informal group norms can develop which are able to take precedence over the management's views on a acceptable level of production.
- Workers which exceeded or did not achieve such levels risked being ostracised by their fellow workers.
- For many workers being a member of the group would be more important than seeking or being given promotion and therefore having to leave the group.
- If management takes an interest in the workforce and allows them some degree of self control and determination, they will feel less alienated to the company
- Where norms of co-operation and high output are established, the physical conditions of the workplace will be less important.

(p. 54)

These findings will to an extent chime with many schools. Although much of a teacher's work is individual, social aspects such as staff room camaraderie can provide support, and help individuals to perform well. And teachers who do not 'pull their weight' can find themselves ostracised. For quite a number of teachers there is little desire or incentive to seek promotion, and good heads do involve staff in decision making. Also many schools have difficult physical conditions, and in spite of this they perform well and staff enjoy high levels of co- operation.

However, many of the reforms to the organisation of the education system do not appear to be 'informed' by the work of Mayo or other human relation theorists. For those working in the system, many of the new practices which have been introduced appear to be aimed at discouraging overt dissent by increasing accountability and using 'harder', less trusting and perhaps less professional employment practices.

Morgan and Allington (2002) in their article *Has the Public Sector Retained its 'Model Employer' status?* discuss two different forms of Human Resource Management (HRM), Hard (*utilitarian instrumentalism*) and Soft (*development humanism)*. They provide evidence that the public sector is moving increasingly towards the hard model, which is antithetical to Mayo's arguments (in effect hard systems follow a rational / scientific model). Examples, they cite include, the increasing trend towards temporary and agency staffing, which do increase the flexibility for the employer, but reduce the ability for employees to develop 'informal group dynamics'. In a similar way the reduced influence and probably consequential membership of the unions has reduced the ability of employees to form such groups. Increasingly, teachers are assessed and held accountable for their own (apparent) performance, rather than their contribution to their department or school, and this

too can increase competition and potentially discourage meaningful co-operation.[58] Various performance indicators are widely used in these hard HRM processes, and this may well for many affect their perceptions and feelings towards these indicators.

However, one key finding from Hawthorne, the importance of feedback (see Parsons 1978 and Fitz-Gibbon 1996) is an important part of the modern education employment policy and practice. Good and effective schools and their heads have of course always given their staff feedback, however this process is now more formally recognised. Regular appraisal is now a legal requirement in schools, and part of the official performance management systems. Although, still much of the formal system tends to be 'hard' HRM, being top-down and is contrary to good industrial practice (Fletcher 1993, Hequet 1994) in that personal performance is tied to the remuneration and accountability processes[59]. Furthermore, there is little evidence of softer, bottom up systems, such as 360 degree or 'up line' appraisal, although perhaps, such approaches are inhibited by 'both sides' because of historical and cultural factors. It is of concern that much of the current government policy would appear to be aiming to increase demarcation; for example, the frequent emphasis on the critical importance of the head, rather then the 'team'. In essence whether a 'hard' or 'soft' model is adopted largely depends on the employer's underlying attitude and philosophy towards the employees; a point considered in the next section.

McGregor: Theory X and Theory Y

Douglas McGregor (1906 – 1964) was a social psychologist, who was particularly concerned with the human aspects of organisations. For him this was the key to an organisations success, and a better society:

> The ingenuity and the perseverance of industrial management in the pursuit of economic ends have changed many scientific and technological dreams into commonplace realities. It is now becoming clear that the application of these same talents to the human side of enterprise will not only enhance substantially these materialistic achievements, but will bring us one step closer to' 'the good society '.
>
> (McGregor 1957a:180)

Leading on from this philosophical standpoint, McGregor is best known for his Theory X and Theory Y which were published in his seminal book, *The Human Side of Enterprise*, (1960). This theory set out two distinct views of which an organisation's management could take of its workers.

Theory X

1. The average human being has an inherent dislike of work and will avoid it if he can. Thus management needs to stress the need for productivity and offer incentive schemes.

---

[58] See for example Williams (1999) and Revell (2000), who discuss two successful schools, Hurlingham and Chelsea, and Thomas Telford, which have introduced a variety of 'hard' HRM techniques.
[59] And potentially the disciplinary process.

2. Because of this human characteristic of a dislike of work, most people must be coerced, controlled, directed and threatened with punishment to get them to perform adequately to meet the organisation's objectives.

3. The average human being prefers to be directed, wishes to avoid responsibility, has relatively little ambition, and above all wants security.

Theory Y

1. The expenditure of physical and mental effort in work is as natural as play or rest. The ordinary person does not inherently dislike work: according to the conditions it may be a source of satisfaction or punishment

2. Workers will exercise self-direction and self control in service of objectives to which he committed. External control is not the only means of obtaining effort.

3. The most significant reward that can be offered in order to obtain commitment is the satisfaction of the individuals self actualising needs. This can be a direct product of effort directed towards organizational objectives.

4. The average human being learns, under proper conditions, not only to accept but to seek responsibility.

5. Many more people are able to contribute creatively to the solutions of organisational problems than do so.

6. At present the potentials of the average person are not being fully utilised.

(Based on Pugh *et al* 1971:149)

These two sets of assumptions take very different positions, with Theory X taking a somewhat cynical view of workers. Control is primarily external, and imposed upon subordinates by their managers. In contrast, Theory Y takes a far more optimistic view of human nature, and emphasises the potential initiative and self-motivation which may be exhibited by individuals. It is however significant to note that McGregor chose the terms X and Y to give an impression of neutrality, and therefore avoiding connotations of 'good' and 'bad'. The table below from Koontz et al (1980) illustrate the sorts of managerial behaviour which is associated with the theory:

Table 3: McGregor's Theory X and Theory Y

	Theory X	**Theory Y**
Selected key managerial activities	**People dislike work; people must be forced to work; people do not willingly assume responsibility**	**People like work; people work best under self-direction; people like to assume responsibility**
(a) Planning (including setting objectives	Superior sets objectives for subordinates Little participation in setting targets and development planning Few alternatives are explored Low commitment to objectives and plans	Superior and subordinates set objectives jointly A great deal of participation in target setting and planning Many alternatives are explored High commitment to objectives and plans

(b)Leading	Leadership is autocratic, based on authority only	Leadership is participative and teamwork is based on competence
	People follow orders, but hidden resistance and mistrust exists	People seek responsibilities, feel accountable and are committed to performance
	Communication is one way, top down with little feedback	Communication is two-way with a great deal of feedback
	Information flow is limited	Necessary information flows freely
(c) Controlling and appraising	Control is external and rigid	Control is internal and based on self control
	Superior act as a judge	Superior acts as coach
	Low trust in appraisals	High trust in appraisals
	Focus is on the past, with an emphasis on fault finding	People lean from the past, but focus on the future; feedforward control emphasizes problem solving

Based on Koontz *et al* (1980, 617)

Perhaps surprisingly McGregor did not intend that his theories would represent the two extremes of a continuum upon which all managers and organisations could be placed. However, as with many organisations, the education system does not easily fit in to either X or Y. In practice, examples of behaviour from both sides can be observed. It is also quite possible for one level of an organisation to embrace one theory, and another level the other; furthermore, there can be a significant difference between espoused views and what actually happens.

In terms of planning and setting targets for example, this should in theory be done by the governing body (QCA 1999) in negotiation with LEAs (ie. theory Y), although, in practice various pressures are brought to bear on schools which limit meaningful negotiation[60]. Furthermore, the trend (Hackett 2000) is towards the government setting targets centrally (ie. theory X), although there is some evidence (Passmore 2002a) in spite of this that teachers are reasonably committed to their objectives (Y).

In a similar way, the actual curriculum and pedagogy are very much in line with theory X, with them being centrally determined, and their being few options for schools to explore any alternatives. Although again it would seem that many schools and teachers are quite happy with this (Earl *et al* 2001, Passmore 2002a). The governments leadership model is very much based on theory X, although, heads in many good and effective schools adopt a theory Y approach with their staff.

However, in terms of control the picture is less ambiguous. Schools are largely controlled by external mechanisms, such as Ofsted and the performance management system (ie. Theory X). Governing bodies do of course play a part, however this is precisely defined and tightly controlled.

[60] LEAs have to ensure that school's targets are challenging, and are required to ensure that their own targets are met.

Within these control systems individual performance appraisal is an important part which McGregor (1957b) described as being based on Theory X:

> Even a cursory examination of conventional programs of performance appraisal within the ranks of management will reveal how completely consistent they are with Theory X. In fact, most such programs tend to treat the individual as though he were a product under inspection on the assembly line.
>
> (p. 179)

This blanket interpretation would seem a bit hard on performance appraisals *per se*. In education they tend to be top-down but can nevertheless provide a useful opportunity for positive two-way communications. Koontz *et al* (1980) suggest that McGregor's concern was mainly to do with managers having to 'play god' in performance appraisals, and that this situation could be improved by using more objective criteria. However, in terms of the actual process in education only a small number of relevant objective criteria are available, much of a teachers performance is dependent (quite rightly) upon many largely intangible factors.

It is difficult to assess whether education is moving towards theory X or Y. In terms of espoused policy this is very much in the direction of Y. The white paper, *Schools Achieving Success* (DfES 2001a) is very much couched in terms of theory Y. Although in practice much appears to be based on theory X, with for example many centralised (and effectively compulsory) initiatives and strategies. A general move towards theory Y would seem to be a good thing. There are compelling arguments that individuals in complex systems like education need to work within such frameworks if they are to be effective, in the broadest sense; for example using their imagination and initiative to address difficult and ambiguous issues. However, it would be wrong to assume that such an approach is necessarily easy and without problems, indeed the demands on the individuals can be far greater, than simply 'following orders'. Drucker describes the experience of Abraham H Maslow, who was the leading theorist on motivation (eg. Maslow's well known *Hierarchy of Needs* 1970) and a strong believer on the value of the theory Y approach (Koontz *et al* 1980):

> Maslow spent one year working closely with a small company in Southern California which at the time tried to practice Theory Y...[Maslow] pointed out that the demand for responsibility and achievement may well exceed what any but the strong and healthy can take. He sharply criticized Theory Y for 'inhumanity' to the weak, the vulnerable, the damaged, who are unable to take on the responsibility and self discipline which Theory Y demands. Even the strong and healthy, Maslow concluded, need the security of order and direction; and the weak need protection against the burden of responsibility.
>
> (Drucker 1999:221)

It is difficult to generalise too much from this, as much depends on the particular situation and needs of the organisation. It does however illustrate the potential difficulties of changing from X to Y. Indeed, if the government were to abandon some of their tightly defined strategies, such as the

NLNS, this might not be viewed by all teachers as a 'wonderfully empowering experience': again much would depend upon the context. The external evaluation (Earl *et al* 2001) of the NLNS found evidence to suggest that schools and teachers welcomed the highly prescribed methods and accompanying materials. This they suggested helped teachers cope with the demands of both these strategies and the many other new initiatives. Whether in the long term this will help develop the skills and abilities of teachers generally, and from this the overall quality of the education system, is another matter.

McGregor (1957a), like a number of other theorists of the time, emphasised the importance of the work situation and context. He pointed out that the behaviour of the workers was largely a reaction to the 'scene set' by the organisation:

> The social scientist does not deny that human behavior in industrial organizations today is approximately what management perceives it to be. He has, in fact, observed it and studied it fairly extensively. But he is pretty sure that this behavior is *not* a consequence of man's inherent nature. It is a consequence rather of the nature of industrial organizations, of management philosophy, policy, and practice.
>
> (p. 175)

In practice there are many examples of employees being very effective in one situation or context, and very ineffective in others; indeed one only has to apply this to children and how they perform in different classes. He went on to argue that many organisations adopted theory X in the mistaken notion of what is cause and what is effect, and that this would lead to 'self-fulfilling prophesies'.

Although, the Hawthorne studies and Theory X and Y have made an important and valuable contribution to the theory of management (Shafritz and Ott 1992, Pugh and Hickson 1989), like all models they do have their limitations. McGregor's work, for example, was carried out in the 1950s and 60s, a time of low unemployment and general optimism in terms of businesses. There was little need to compete with the tiger or pacific rim economies, and even education in those days appeared a far 'simpler' process. Consequently, such models are less able to deal with the many contradictions and paradoxes of present day management, and in particular the complexities of public management with its many different influences and objectives. Others such as Morse and Lorsch (1970) *Beyond Theory Y*, and Ouchi (1981) *Theory Z: How American Business Can Meet the Japanese Challenge*, have further developed the theories to meet the needs of such organisations. Unfortunately, there is insufficient space to look at some of these, and therefore the next section 'jumps' to *Hybrid* management theory.

Hybrid School

The term Hybrid is used here to describe the 'modern' or contemporary approach to management. In many respects these combine the theories from both the rational / scientific and the behaviourist school. For example, one of the approaches considered here, Business Process Re-engineering, has elements of scientific management in that actions are carefully planned and theory Y in that stakeholder involvement is encouraged.

A substantial number of Hybrid approaches have been developed over the last thirty or so years, indeed many have been described as 'Fads' (Caulkin 1996, Shapiro 1998). A number though have proved to be more enduring, for example, the 'Learning Organisation' philosophy.

This section firstly looks briefly at *Management By Objectives* (MBO), which helped set the scene for hybrid management, and the well known and influential work of Peters and Waterman (1982), relating to 'Excellent Organisations'. Following this, three different hybrid approaches: *Total Quality Management*, *Learning Organisations*, and *Business Process Reengineering*, are considered in more detail.

The background the Hybrid approaches

The impetus for the hybrid approaches was caused by two factors in particular; the general world economic recession of the 1970, in part due to the OPEC 'oil crisis', and the relative economic success of Japan. This resulted in a general tightening of economic controls, which was particularly evident in the public sector (see chapter 5) and the development of new management practices to meet these challenges.

Much of Japan's success at this time, has been attributed to the different management culture which existed in Japan, and in many respects was more in tune the 'behaviourists' rather than the bureaucratic structures of the west. Konosuke Marsushita the founder and Executive Director of Matsushita Electric, (one of the largest electronics companies in the world), described these different approaches;

> We are going to win and the industrial West is going to lose out: there's nothing much you can do about it, because the reasons for failure are within yourself...for you, the essence of management is getting the ideas out of the heads of the bosses into the hands of labour...for us, the core of management is precisely the act of mobilizing and pulling together the intellectual resources of all employees...only by drawing on the combined brainpower of all its employees can a firm face up to the turbulence and constraints of today's environment.
>
> (Matsushita 1982)

It would of course be wrong to tar all the western companies with this broad brush, there are many successful western companies which have a long history of such an approach to management: for example, Hewlett and Packard and Proctor and Gamble (Pugh and Hickson 1989). Many theorists (eg. Ouchi 1981 and Deming 1986) largely support the views of Matsushita, and have concluded that these cultural differences led to the success of Japan during the latter half of the 20th centaury[61]. Principles based on these difference have led to the formation (and language) of modern hybrid management theory.

Management by Objectives

Peter Drucker the highly prolific theorist and writer is largely credited (Shafritz and Ott 1991) with identifying and formalising the principles of Management By Objectives (MBO) in his book *The Practice of Management* (Drucker 1954). Although the system became very popular during the 1960's and 70's interest began to wane during the 1980's. The system is quite simple; objectives are set by the management for individuals and functional areas in the organisation, and the role of the subordinates is to meet these objectives (they may have some discretion in how this is achieved). Although MBO is now largely discredited as a management system[62] (Koontz *et al* 1980, Shafritz and Ott 1991), it is relevant to both the development of hybrid management theory and current day public sector management.

As a system it combines the hard management principle of setting objective targets, with the potentially softer principle of allowing individuals some degree of freedom and creativity in meeting their targets. There is no particular requirement as to what the targets are based on. They may for example be key output indicators or more specific process indicators. Within education the current Performance Management Framework (DfEE 2000a), has many similarities with MBO, objectives are firstly set, then monitored through the time period (typically a year) and in due course reviewed;

> **Planning:** team leaders discuss and record priorities and objectives with each of the teachers in their team. They discuss how progress will be monitored.
>
> **Monitoring:** the teacher and team leader keep progress under review throughout the cycle, taking any supportive action needed.
>
> **Review:** the teacher and the team leader review achievements over the year and evaluate the teacher's overall performance taking account of progress against objectives.
>
> (DfES 2001f: 5)

---

[61] There is some debate that this Japanese approach has now led to the economic difficulties currently evident in Japan and other Pacific Rim countries.
[62] Drucker's more recent writing (eg Drucker 1980) makes no mention of it.

The main complaint against MBO is that it encourages dysfunctional behaviour. Below is a summary by Koontz *et al* (1980) of some of the widely accepted problems of MBO systems:

- encouraging short-term behaviour,
- difficulties setting the objectives at the right level (ie too hard or too easy),
- encouraging inflexibility,
- using numbers for objectives which can not be effectively quantified, and
- attempting to apply the same numerical objectives to different units of the organisation which operate in significantly different circumstances.

Koontz *et al* (1980: 209)

And in the guise of performance management systems for schools, Cowie (2000) argues that:

> Because performance management promotes the cult of performativity, it may stifle creativity, inhibit improvement, distort educational purposes and encourage an emphasis on the readily testable at the expense of more diffuse aims, discouraging qualitative monitoring and work against schools being responsive to the communities they serve. Instead, performance management decentralizes in order to tighten control, giving rise to a version of accountability that *"reinforces conventional hierarchical control at the level of the microstructure* "(Mintzberg, 1996, p. 81).
>
> (p. 2)

In spite of the potential dysfunctional effects of MBO, objectives or targets do nevertheless play and important and often valuable part in managing organisations. For instance, in terms of individual appraisal they can provide clear unambiguous numerical data, upon which potentially fair judgements can be made. Furthermore, as Koontz *et al* (1980) point out using MBO systems can increase commitment and involvement from individuals when they are able to choose and develop their own methods. Indeed, Drucker (1977) although no longer mentioning MBO, outlines many important issues relating to the use of objectives in general, and specifically points to the potential danger that, "...objectives that become straitjackets do harm". (p. 93). He concludes his arguments by stating that objectives should provide, direction and commitment directed towards the future of the organisation. The key to all these points is selecting appropriate objectives which reflect the values of the organisation, and then using them in an appropriate and supportive way.

Excellence

The idea of excellent organisations came to particular prominence in Peters and Waterman's classic and best selling business book, *In search of excellence* (1982). This work has had a major impact on both business and public sector management (Pollitt 1993, Hughes 1994, Stacey 1993). It was written at a time of the increasing awareness of the need for all companies to improve their performance particularly in terms of their customers. Pollitt (1993) points out that this increasing culture of 'excellence' started in the US in 1980, and was then adopted in varying degrees in the UK.

Currently, the 'excellent' word appears in many policies and policy documents: for example, '*Excellence in Schools* and *Excellence in Cities*, as well as publications such as, '*The Pursuit of Excellence*', (Barber 1999), *The Business Excellence model*,[63] (Ofsted 1998), '*Achieving Excellence through the National Curriculum*', (Blunkett 2000); indeed a keyword search of the word 'excellent' on the DfES website brought up nearly every document that has been recently written!

Peters and Waterman's work was based on a study of 43 well known, and what were commonly regarded as being very successful American corporations. Although these were all from the fortune 500 list, the argument was that *any* organisation, anywhere in the world, could imitate these features, and in due course become 'excellent' and successful. Peters and Waterman identified eight recurrent features of these organisations:

1. **Stick to the knitting.** Organisations should stick with their core strengths, which the managers knew and understood best.

2. **Close to the customer.** The excellent companies by listening to their customers became better adapted to their environment than their competitors.

3. **Productivity through people.** There was a positive approach to generating a feeling of belonging on the part of the employees, and a corresponding improvement in performance.

4. **Autonomy and entrepreneurship.** This encouraged individual initiative and to a degree, risk taking.

5. **Hands on, value driven.** The leaders of the organisation define the vision and core values, and instil a common culture.

6. **Bias for action.** The emphasis here, is against too much lengthy analysis and planning, but rather one of trying it and seeing.

7. **Simple form, lean staff.** Developing flatter organisations with authority pushed down the line. Rigid job descriptions are discouraged.

8. **Simultaneous loose - tight properties.** Excellent companies were found to be able to allow, simple fluid structures of autonomy and individual empowerment (loose), with short term financial controls and performance evaluation (tight). In effect the companies were both decentralised and centralised.

(Based on Peters and Waterman 1982)

A number of parallels can be identified with the development of education and the broader public sector. For example, the development of governing bodies and the involvement of parents in Ofsted inspections has encouraged schools to become much closer to parents and the wider community (2). The performance management system aims to recognise the importance and encourages the development of individuals working within the education system (3). There is some evidence of

---

[63] A definition *they* use to describe the inspection framework.

schools becoming flatter and leaner organisations, although this many not necessarily be from the point of view of developing 'excellence' but rather meeting financial targets (7). The idea of *loose – tight* properties relates closely to the principle of a 'hard' systems approach which is adopted by the government, combined with the 'softer' management techniques which are supposed to be 'taught' to school leaders (NCSL 2002).

As previously mentioned in chapter 3, Ouston (1999) highlight the similarity and overlap with a list of the 'features of effective schools' by Barber *et al* 1995:

- professional leadership: strong, purposeful, involved
- shared vision and goals
- a learning environment
- concentration on teaching and learning
- explicit high expectations
- positive reinforcement
- monitoring progress
- pupil rights and responsibilities
- purposeful teaching
- a learning organisation
- home-school partnerships

Stacey (1993) points out that like so many recipes for success, the Peters and Waterman study did not prove to be 'the definitive route'. Indeed, since their study, quite a number, two thirds within five years, of the organisations in the study have 'fallen' some way from their peak. (for example Wang, Atari and even IBM). This is however not to suggest that the study does not have some value, and of course a simple explanation might be that their competitors have read the book and also found success!

Although much of what Peters and Waterman said was not original, it was the bringing together through their research of these ideas and their very accessible writing style, which very much contributed to the phenomenal success of the book. There is no doubt that their study occurred at the right time and encouraged many business leaders to examine their approaches, and to become receptive to the 'new' hybrid management approaches.

Total Quality Management

The term Total Quality Management (TQM) refers to quality programs and systems, which were mainly developed during the 1980's and 90's. These have the general aim of improving the quality of products or services, which would in turn lead to better business performance. Although, the systems were originally developed for industry, initially as a means of meeting the competitive challenges of the Japanese industries, they have also been found to be applicable to service and public sector organisations (Deming 1986, Murgatroyd and Morgan 1993). In education for

example, Ofsted claim that their inspection framework is: "virtually indistinguishable in overall shape from the structure of a 'Business Excellence model' which is based on TQM" (Ofsted 1998 p. Annex A). Furthermore, the government in their white paper, *Modernising Government* (Cabinet Office 1999) advocate the adoption of the European Foundation for Quality Management (EFQM) Business Excellence Model (BEM) for the public sector generally, and again this is a TQM model.

There is no such thing as *the* TQM system, but rather a number of different systems, which have been developed, for example by; Philip Crosby, Joseph Juran and W. Edward Deming, or indeed by organisations themselves. In general the various systems have the following features;

- Focus on the needs of the market (or customers)
- Achieve top quality performance in all areas of the organisation
- Develop measure of performance
- Establish simple procedures for ensuring quality performance
- Continually review processes to eliminate waste
- Ensure effective communication
- Seek never-ending improvement

(Based on Freeman 1993)

As well as differences in the various systems, there is also variation in how they should be applied. Some writers advocate an approach which concentrates on specific areas of the organisation. For example, Freeman (1993) and Sallis (1993) argue that the key to an organisation's success is the pursuit of ever improving quality from the perspective of the customer (or student). Others such as Oakland (1995) and Deming (1986) emphasise the importance of an overall or cultural approach to quality.

The model developed by W. Edwards Deming is considered below, although he did not in fact refer to it as TQM, but rather, 'The Deming System of Profound Knowledge'™. Deming's work has received widespread approval, and is considered by many (for example, Crawford and Shutler 1999, Weller 1995), as being particularly relevant to the education system. This system originates from Deming's work in the 1950's in Japan, where he was credited with contributing to the 'Japanese industrial miracle' (Deming Organisation 2000).

Deming's system consists of '14 points for management', which he claimed are applicable to all organisations, whatever their size and type:

1. **Create constancy of purpose for improvement of the product and service, with the aim to become competitive and to stay in business, and to provide jobs.** This emphasises the need for long term views, and being able to meet the constantly changing needs of their customers.

2. **Adopt the new philosophy.** Deming makes the assumption that the 'old philosophy' accepted faults such as delays, mistakes and allows the use of poor materials; and that to be successful the managers must accept a 'new philosophy'.

3. **Cease dependence on mass inspection to achieve quality.** This moves the primary quality inspection down to those carrying out the production. Deming emphasises the need for training and the development of techniques for people to do this.

4. **End the practice of awarding business on the basis of price.** Many expensive and long term mistakes have been made by buying the cheapest. Deming also emphasises the importance of developing long term supplier relationships, which enable and can encourage the customer and supplier to work more closely.

5. **Improve constantly and forever the system of production and service, to improve quality and productivity, and thus to constantly decrease costs.** This has now become commonly known as 'Continuous Improvement'.

6. **Institute training on the job.** The greatest waste in any organisation is the failure to use and develop the talents of its people. There can be many costs and risks employing a new specialist for a particular function, which an existing member of staff, with the appropriate training may be able to carry out, to the benefit of both the individual and the organisation.

7. **Institute leadership.** This requires a new kind of management, one which has moved away from a command and control approach, to one based on an inter-personal leadership approach.

8. **Drive out fear, so that everyone may work effectively for the company.** The principle here is that staff motivation is largely dependent on security. Deming believes that people want to do a good job, assuming they work in an environment which encourages them.

9. **Break down the barriers between departments.** This aims to encourage people to work together as a team throughout the organisation, which will reduce dysfunctional competition between individuals, and groups of workers.

10. **Eliminate slogans, exhortations, and targets asking for new levels of productivity.** Deming was concerned that these lead to an dysfunctional adversarial approach to management. In addition most of the problems to do with production are systemic, and beyond the control of the workers.

11. **Eliminate work standards that prescribe numerical quotas.** Deming argues that working to numerical quotas often leads to cutting corners and an overall reduction in quality of the 'whole' product.

12. **Remove the barriers that rob people of their right to pride of workmanship.** Deming argues against such things as appraisal systems and merit ratings, which he believes puts staff in competition with each other, and can discourage teamwork.

13. **Institute a vigorous programme of education and self-improvement.** This is aimed at all individuals within the organisation.

14. **Put everyone in the company to work to accomplish the transformation.** It is essential to involve everyone in the transformation to a 'new' quality culture.

(Based on Deming 1986)

Deming also identified barriers (the seven 'deadly diseases') to the 'new management'. Most relate to ownership and financial structures of commercial organisations; however relevant to this thesis.

are the widespread use (or mis-use) of performance appraisals, and high stakes external indicators, ie KPIs.

Deming argued against prescribed numerical quotas (11) and the ratcheting up of targets (10) which are at the heart of the government approach to 'raising standards'. The current systems of accountability relies heavily on mass inspections (3), ie by Ofsted, and the government's approach of 'name and shame' would seem to 'drive in fear' (8). There are examples of the government encouraging professional on the job training (6) although this is often in response to teacher shortages. There is also evidence of schools and teachers adopting 'the' new philosophy, although the government interpretation of this would seem to be narrowly defined as driving up KPIs.

Much of the attraction of TQM for many schools and educational institutes is that it highlights the importance of quality, and on the face of it a new more enlightened approach to management. Bonstingl (1993: 66) points out: "TQM provides a welcome point of departure from this outmoded, (Tayloresque) counterproductive model of education"....and suggests the power of TQM lies in its philosophy of continuous improvement, or a 'yearning for learning'. Sallis (1993) highlights the integrated nature and potential power and value of TQM but also points out that it is not a magical overnight cure;

> It is not just another initiative. It is a philosophy and a methodology which assists institutions to manage change, and set their own agenda for the plethora of new external pressures.... In the industrial sphere it is seen as the means by which beleaguered economies in the West can transform themselves to compete better with the fast growth economies of the Pacific rim. There are those in education who believe that TQM properly applied to it can complete a similar transformation. However, TQM does not and will not bring results overnight. The essence of TQM is a change of culture. Changing the culture of an institution is a slow process, and one that it is best not rushed.
>
> (p. 99)

Quong and Walker (1996) describe how TQM can help in changing and restructuring a school, which has to meet the potentially contradictory needs of, "[averting] a fiscal crisis and its community's demand for an organisation that was more caring and responsive to student, parent and teacher needs" (p. 219). They found that the process was far from simple, it required a lot of hard work and commitment from all parties, and was most certainly not a quick fix. However, in the long term they felt it had produced many tangible benefits. The issue and indeed conflict of time scale is an important one. Currently, like businesses, schools are measured on an annual basis, and initiatives which do not provide returns within that time frame may well be unattractive to policy makers. Crawford and Shutler (1999) point out that using Crosby's 'zero defects' model of TQM is more appropriate to 'quick returns'. Conversely, Deming's approach takes a longer term systemic view of an organisation and its performance. Therefore, this model would seem to be at odds with aspects of government policy such as the PSA targets for 2002 (Treasury 2000).

Chappell (1994) points to the considerable variations in the application of TQM and in general concludes that they help improve the quality of the product (ie the teaching), but not so much the quality of the organisation. This would suggest that the more holistic application of TQM is loosing out to an approach which improves short term tangible results. Within higher education, Holmes and McElwee (1995) point out that the 'modern' performance culture and accountability structures is antithetical to the underlying philosophy of TQM;

> Quality management has, in the university and indeed in further education become a 'strategic issue'. However, unlike any other sector in the economy, TQM has not been given the opportunity to be used as a strategic driver - not because of lack of interest, but because the imposition of external control mechanisms have prevented HEOs dealing with issues of quality enhancement. As regards quality, HEOs have in the last four years been simply trying to come to terms with dealing with the interests of the Higher Education Funding Council and the Higher Education Quality Council at the expense of quality enhancement.
>
> (p. 8)

In essence TQM systems do have much to offer, however their potential long term benefit, will not necessarily provide short-term KPI improvement.

The Learning Organisation

Organisations that learn have of course existed for hundreds or thousands of years, however it is only more recently, since the 1980's in particular, that the term 'learning organisation' has gained common usage in management theory. Two books in particular helped set the scene; *The Fifth Discipline: The Art and Practice of the Learning Organisation* (Senge 1990) and *The Learning Company: a strategy for sustainable development* (Pedler *et al* 1991). Many of the principles build on other work, such as that of Argyris and Schön's (1978) who looked at different types and levels of learning. Indeed, the importance of learning has been highlighted by many others, for example. Peters and Waterman (1982) succinctly stated that, "The excellent companies are learning organisations".

There is no simple definition of what a learning organisations is, indeed the there are many definitions and interpretations, with much depending on the particular context. However, to help set the scene, the following quotes aim to illustrate some of the fundamental principles:

> Organisations where people continually expand their capacity to create the results they truly desire, where new expansive patterns of thinking are nurtured, where collective aspiration is set free, and where people are continually learning how to learn together. (Senge 1990)
>
> A Learning Company is an organisation that facilitates the learning of all its members and continuously transforms itself. (Pedler *et al* 1991)

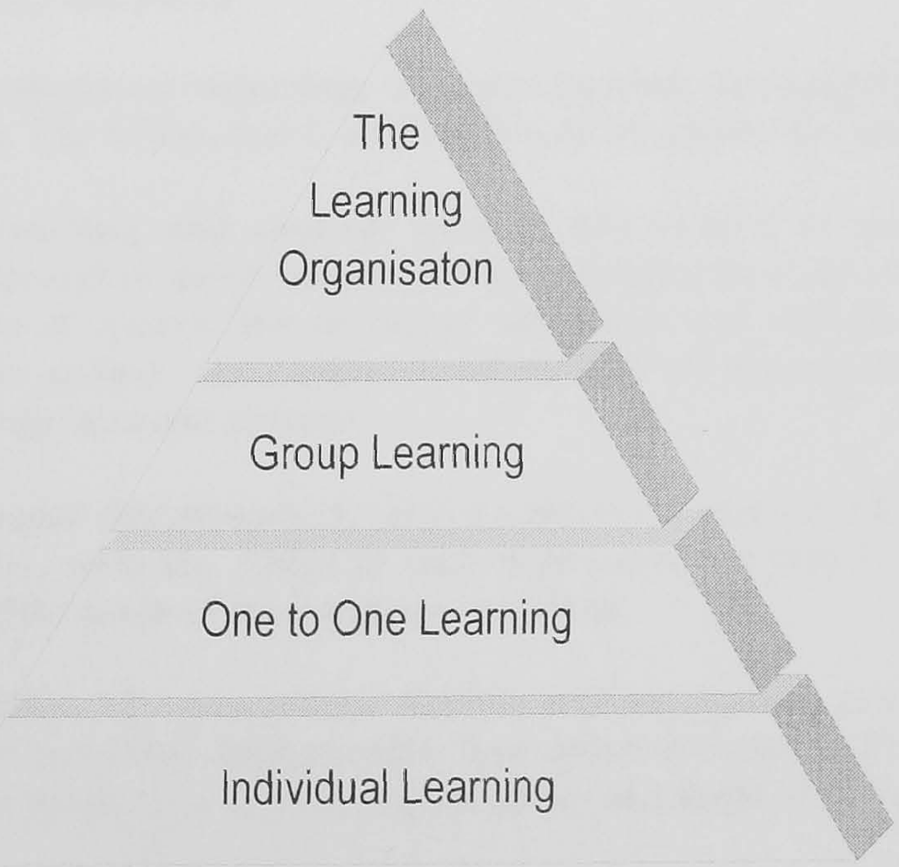
> Creating an environment where the behaviours and practices involved in continuous development are actively encouraged. (Mumford 1995)

These definitions can be expanded to give six key principles of learning organisations:

- Encourage innovation and experimentation
- Have a relatively blame free culture
- Encourage participation in policy and strategy formation
- Ensure effective internal and external communication and co-operation
- Create supportive information systems
- Encourage a long term and holistic view of the organisation

A common theme from these is the importance of a climate that allows learning to take place at *all* levels of the organisation. Mumford (1995) argues that learning starts at the individual level, and this can then progress throughout the organisation, to give a 'Learning Organisations'.

Figure 6: A learning organisation model



From Mumford, A (1995) The Learning organisation in Review

As previously discussed in chapter 2, what constitutes *the* organisation can be unclear. For example, does the education system consist of a number of autonomous self-managing schools, or is it a cohesive national system, where individual schools are in effect 'sub-branches'? Principles such as 'Local Management of Schools', suggest that schools should be discrete organisations, and

therefore able to learn and develop on an individual basis. On the other hand, the organisation and control structures which have been imposed as a result of the 'New Educational Management' (see chapter 7), suggest that education is far more of a unified or centrally managed system, and therefore the whole system has to learn, with the 'learning' being passed down the system.

Many educational writers and theorists have extolled the virtues of 'learning organisations' (for example, O'Sullivan 1997, Diggins 1997, Gunter 1996). Intuitively, the concept is very attractive or indeed a natural one for educationalists. To help define in more detail what constitutes a 'learning organisation' Pedler *et al* 11 point framework is given below, which has been widely applied to both industrial and service organisations:

1. **Learning approach to strategy:** Business plans are evolved and continually modified as necessary. Small scale experiments with feedback and review can be built in to the process, to allow continual improvement.

2. **Participative policy-making:** All members of the company take part in policy and strategy formation, also stakeholders; employees, customers and suppliers, as well as the management, influence the policy making. Appraisal and career planning decisions can contribute to the company's strategy and policy.

3. **Informating:** Information technology is used to increase understanding of what is going on in the organisation. The information is not simply used for reward and punishment.

4. **Formative accounting and control:** Ensures that systems of accounting, budgeting and control are structured to assist learning. This encourages an ethos of self-responsibility, with the development of systems that encourage individuals and units to act as 'small businesses' within the larger system. Accountants or others with a similar controlling function are also perceived as consultants and advisers.

5. **Internal exchange:** Departments develop a customer / supplier relationship with each other, and agree quality, costs etc. Although each department will seek to 'delight' others, they do remain aware of the needs of the company as a whole.

6. **Reward flexibility:** This can include flexible working patterns and alternative non-financial rewards, as well as looking fundamentally how different people and roles are rewarded. All of the company are involved in determining the nature and shape of the rewards system.

7. **Enabling structures:** Roles are loosely structured in line with the needs of internal customers and suppliers, and to allow individual growth and development. Rules and procedures can be changed after review and discussion, and boundaries can flex in response to changes.

8. **Boundary workers as environmental scanners:** Individuals who have contact with the external environment, eg. customers, suppliers, competitors and associated businesses, are used as sources of information which will be used by the organisation.

9. **Inter-company learning:** This can take the form of 'benchmarking', where a company learns from other companies in *different* situations and industries[64].

---

[64] This is different to the official educational benchmarking system. The important point being that an organisation can learn from other industries, as well as other organisations in their own industry.

10. **Learning climate:** The primary role for managers is to facilitating employees experimentation and learning, in a climate that allows mistakes (within reason) to be made. In addition it discourages the principle of certain procedures and practices of being sacrosanct

11. **Self-development opportunities for all:** Resources and facilities for self-development are made available to all employees. A wide range of training and development materials and resources are available, and are offered to meet individual needs and aspirations. The individual has a substantial influence on the training they receive.

(Based on Pedler *et al* 1991)

The principles of informating (3) and formative accounting (4) are particularly relevant. A significant amount of performance information is used by schools formatively, however some is also used for 'punishment and reward'. In addition, senior staff or LEA personnel, can adopt a consulting or advisory function, which may encourage 'learning' to take place at the school level. In many schools a high degree of internal exchange (5) exists with departments helping each other, however there is a danger of accountability systems encouraging a 'blame culture', and therefore reducing co-operation. Some degree of inter-company learning (benchmarking) can occur, although this is rarely with other businesses. The planning process can enable a degree of participative policy making (2). However, in terms of other key areas such as enabling structures (7) and a learning climate (10) schools do not score well.

Although few would disagree that learning organisations and a learning culture is a good thing, even if it works well there are still potential problems. Nutley and Davies (2000b) point to the problem in the health sector of individual organisations learning 'too much':

> Developing learning capacity may lead to more flexible healthcare services and may enable providers and health authorities to meet parts of the government's quality agenda. However, there is no guarantee that learning will lead healthcare organisations in predictable directions. Indeed, the growth of capable and reflective organisations may highlight dissonance between what organisations perceive as appropriate goals (and the means of achieving them) and the directions stipulated by national policy or overseeing bodies. Managing these conflicts will require care.
>
> (p. 1001)

There would seem little doubt that the broad principles enshrined in the learning organisation approach will be the key to a more responsive and better performing education system, however, the key problem is enabling and managing this process within the broader national and political contexts.

Business Process Re-Engineering

The final system considered in this chapter is Business Process Re-Engineering (BPR), which is a 'pro-active' form of organisational development designed to 'drive up' performance. On the face of it this approach seems relevant to the current educational development. Michael Barber, whose thinking and proposals are very much at the heart of the educational reforms, makes quite extensive

reference to reengineering the education system (eg. Barber 1996 and 1997). Davies (1997) too points out that much of Barber's thinking is based on the BPR principles, and states that:

> The era in which we are living demands rapid and fundamental change such that traditional incremental school improvement strategies may be inadequate to meet that challenge. In such circumstances, the necessary paradigm shift for educational leaders may be achieved by using the concepts and approaches from the reengineering movement.
>
> (p. 184)

BPR contrasts with the 'softer' learning organisation approaches, which encourage organic growth from within the organisation, by being far 'harder' and centrally or externally driven. The term was popularised[65] by Michael Hammer and James Champy in their seminal book, *Reengineering the corporation* (1993). In essence they describe BPR as:

> ...the fundamental rethinking and radical redesign of business processes to achieve dramatic improvements in critical, contemporary measures of performance, such as cost, quality, service, and speed.
>
> (p. 32)

As Halachmi and Bovaird (1997) point out BPR is much more than making small incremental changes to existing policy;

> BPR is about breaking off from and doing away with past administrative traditions when marginal adjustments to past practices do not seem to help the organisation in dealing with its current situation.
>
> (p. 228)

Indeed, much of this resonates with the current government's public sector reforms, for example Blair's New Year's message;

> The public sector is undergoing an unprecedented amount of reform to revive it, set it free, from unnecessary constraints. And that reform is often unsettling but I believe out of it will come a new sense of public service ethos where staff and consumers are proud of what our public services can deliver.
>
> (Ahmed 2001)

Hammer and Champy (1993) argued that the design of workflow in most large corporations was based on assumptions about technology, people, and organizational goals that were no longer valid. Others (eg. De Cock and Hipkin 1997) have pointed out that BPR 'took over' where TQM left off, because TQM is not radical enough to meet the highly competitive challenges of the 1990's. Hammer and Champy put forward seven principles of reengineering aimed at streamlining the processes, and which would give significant levels of improvement in terms of quality, time and cost.

---

[65] De Cock and Hipkin (1997) described it as 'hottest' topic in the business press during the 1990s.

1. Organize around outcomes, not tasks.
2. Identify all the processes in an organization and prioritize them.
3. Integrate information processing work into the real work that produces the information.
4. Treat geographically dispersed resources as though they were centralized.
5. Link parallel activities in the workflow instead of just integrating their results.
6. Put the decision point where the work is performed, and build control into the process.
7. Capture information once and at the source.

(Hammer and Champy 1993)

Although the system was originally designed with manufacturing processes in mind, a number of these principles are relevant to service industries and education. For example, the current government's mantra of *'what counts is what works'* indicates a shift towards organising the education on the basis of outcomes (1), or more specifically those measured by KPIs. There is also much evidence of prioritising the processes within the systems, in particular those which contribute to the PSA targets (2). Also developments in information technology have allowed resources to be shared on a wide area basis, for example, the proposals to use 'broadband' based communications to share lessons between schools (4). Claims have been made to localise decision making (6), for example, LMS, although in practice much depends on the particular decisions. The principle of capturing information at source is relevant to key stage testing. A BPR approach would suggest that this be done by the teacher in the class, (as in Scottish primary schools) and not as part of a centralised testing process.

Harvey (1996) identified eight critical success factors for the implementation of BPR, of which two are particularly relevant to the issues considered here. He suggest that the successful implementation of BPR requires the setting of goals which stretch performance, ie "ambitious targets to meet or exceed current world-class performance" (p. 61), and that there should be a high degree of stakeholder involvement. Both of these points chime with current governments education policy.

In practice writers such as Davies (1997) and Fidler (1996)[66] give examples of the application of BPR, such as the teaching of information technology and increasing the number of places in popular schools. However, it is the macro aspects of education policy which are most relevant. Barber (1997) for example looks as the issue of LEAs and considers how they might be replaced by

---

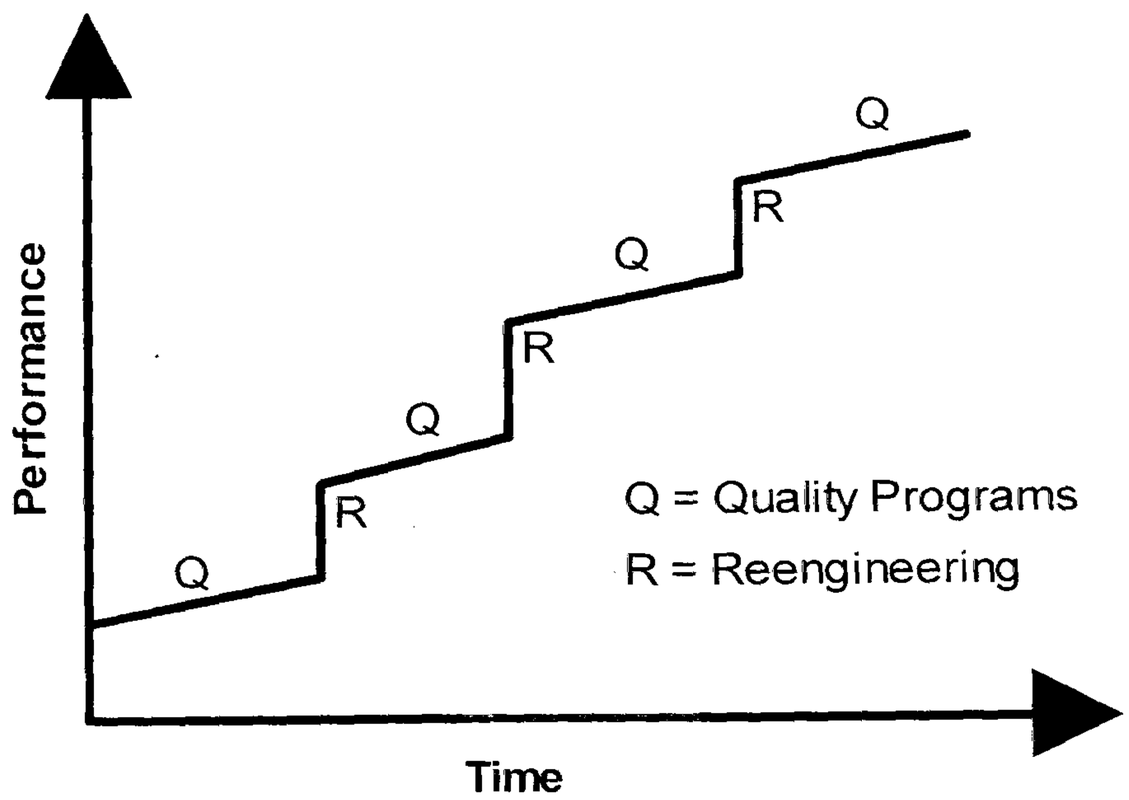
[66] Fidler refers to BPR as Basic Process Reengineering

regional assemblies. And Hargreaves (1994) uses a reengineering approach to look at different staffing structures; ie a smaller core staff on five year contracts which are supported by assistants.

Much of the thinking at this level is inextricably linked to the political processes and ideologies, and therefore is in danger of moving outside of the sphere of what BPR is supposed to do. At the heart of BPR is the need for a 'blank sheet of paper' approach. However, many of the current policy debates are highly coloured by pre-existing views or prejudices, and it would seem doubtful that the necessary uninhibited thinking and policy formulation could be achieved[67].

The rate and pace of change provides a interesting contrast between BPR and other hybrid methods. The key to BPR is step change, which will produce immediate results, whereas other methods such as TQM aim to produce slow steady change. Hammer illustrates this feature of BPR in *Figure 7*:

Figure 7: Quality programmes and reengineering



Hammer 1996: 83

The need for step change in education has been widely promoted by the government. For example, Blair (2001) makes specific reference to the need for 'step changes at key stage three'. And Barber (2000b) states: "The determination of the Blair government to pursue education reform and bring about a step change in the performance of the education service is not in doubt." (p. 3). This would however seem to contradict his previous claim (Barber 1997) which advocated that: "Slow and

[67] There are of course 'blue skies' thinking departments within the government, for example the Forward Strategy Unit.

steady growth [is] better than fast fluctuating growth". (p. 192). And other areas of government policy, for example, the white paper, *Modernising Government* (Cabinet Office 1999) advocate (slow) 'continuous improvement' approaches.

Carr and Johnson (1996) *Table 4* highlight some of the other significant differences between TQM (including the learning organisation principles) and BPR;

Table 4: A comparison between TQM and Reengineering

FACTORS	TQM	REENGINEERING (BPR)
Type of change	Evolutionary – a better way to compete	Revolutionary – a new way of doing business
Method	Adds value to existing processes	Challenges process fundamentals and their very existence
Scope	Encompasses whole organisation	Focuses on core business processes
Role of technology	Traditional support, eg management information system	Use as enabler

Carr and Johansson (1996:16)

Great claims about the benefits of BPR have been made, and in its 'heyday' it enjoyed phenomenal growth, with some 59% of British organisations planning or undertaking some form of such activity (De Cock and Hipkin 1997). However, there have also been claims of a high (70-80%) failure rate of projects (Devine 1994, Doherty and Mistry 1996). Vakola (2000) identifies some of the reasons as: lack of sustained management commitment, unrealistic expectations, and resistance to change. In many cases the benefits have not met the expectations, (or at least those claimed by those selling the consultancy). BPR programmes frequently result in a reduced workforce, both in terms of number and status, indeed Bartholomew (1999) refers to BPR as a fancy word for layoffs.

Although, claims have been made that BPR is a revolutionary approach (Hammer and Champy 1993), other such as Grint (1994) argue that many of the claims are not new: "few, if any, are actually innovations, least of all radical innovations that would support the hype," (p. 191), and Shapiro (1995), simply described the process as just another name for; "run of the mill 'reorg'." (p. 185).

There are many examples of the successful implementation of BPR, although the degree of success is largely determined by the nature of the organisation. Vasey (1994) looked at the effect of BPR on the HMSO, and concluded that there were negative consequences in terms of synergy in large

diversified organisations. This would seem relevant to education, which is a very large and diverse organisation, and where synergy is important at all levels.

Central to BPR is the need to assess value, and for some organisations assessing the value of outcomes might appear to be relatively easy, ie profit or KPIs. However, Halachmi and Bovaird (1997) warn of the dangers in complex organisations of taking a too narrow or simplistic view:

> Public sector managers should use the widest possible definition of 'value' when analysing value-added in process reengineering and should be especially sensitive to the way in which 'value' in the public sector is differently interpreted by major stakeholders.
>
> (p. 234)

This is very pertinent to education. Furthermore, they warn against claiming too much from BPR programmes; "...public sector agencies would be well advised to be conservative in estimating gains from BPR". (p. 227). Not a view that would appear to be shared by the government, (see for example the claims made by Barber 2000a).

De Cock and Hipkin (1997) point out that the most impressive gains are made by companies which are already performing well, and that BPR can have a very negative impact on companies which are struggling. This suggest that BPR may not help under performing schools, or reduce the current performance gap, between 'best' and 'worst'.

Hammer and Champy (1993) recognise that BPR is a very disruptive process, and to avoid undue stress and loss of overall performance, they advocate that the process should be carried out quickly and incisively, and typically within a year. However, as mentioned before, much of the current government's education policy advocates a longer term continuous improvement (ie. a Learning Organisation approach); see for example, the *Leadership development framework* (NCSL 2002). which makes extensive reference to schools developing as learning communities.

This inconsistency has been recognised by Tooley (1999) with his description of 'two Barbers'. For example, Tooley points to the conflict in Barber's thinking in terms of the reforms being a "one off reengineering project or [them] being a process which will *continually* generate novelty and adaptability" (p. 36). The former is BPR and the latter 'Learning Organisation'. Indeed, Barber's (2000b) more recent claim that: "in a democracy – a long term strategy will only succeed if it delivers short-term results" (p. 5); is somewhat perplexing, and would seem to suggest that 'some' BPR is necessary for long term quality development[68].

---

[68] This seems to take us back to Shapiro's 'run of the mill – reorg', rather than 'true' BPR.

In essence it would seem that the government are attracted by the ideas of BPR (especially the language), but in reality realise that developments to complex systems such as education need to be carried out in a more careful and less hazardous way. And more specifically achieving (or being able to claim) step change improvements in KPIs, year on year, is at the very least problematical, and can not carry on indefinitely.

Conclusion

In many respects this chapter has gone full circle, starting with rational / scientific management theory and arriving at business process reengineering; both sharing many common properties. As Halachmi and Bovaird (1997) point out there is a certain attraction of this more familiar and secure form of management:

> BPR aims to put rationality and systematic thought back into management transformation efforts, rather than relying on vaguer notions of inspirational leadership and culture change.
>
> (p. 227)

Many of theories considered in this chapter have some relevance to education, indeed it is not possible to precisely identify one as being *the* education management model. It may be that the new school of educational management studies (EMS), will develop new conceptual models which will be able to do this, although they will have to deal with the tricky political context.

The main problem with trying to identify and describe one particular model of management for education is the many contradictions and inconsistencies in both policy and practice. For example, in one situation or audience, hard BPR step change is advocated to 'drive up' standards, (which are narrowly defined – KPIs). Then at other times, or for other audiences, far softer and more conciliatory messages of partnerships, and the need for continuous development from within the system, are proposed. In a similar way in terms of practice, there are frameworks for the softer TQM or learning organisation approaches, however in reality the actual practice is imposed from above.

Therefore, (at the present time) although there is some evidence of 'softer' management thinking, which recognise the complexities and subtleties of the education; in essence education management is predominately a 'hard' form of bureaucratic scientific management.

Given the political realities of managing the education system, such a view is perhaps not very surprising, and may not necessarily be all bad. Indeed, Jaques (1991) argues that the process of

constantly searching for something better may in itself be counterproductive because it avoids the key underlying issues;

> No amount of exhortation, attitudinal engineering, incentive planning, or even leadership will have any permanent effect unless we understand what hierarchy is and why and how it works. We need to stop casting about fruitlessly for organizational Holy Grails and settle down to the hard work of putting our managerial hierarchies in order.
>
> (p. 262)

The importance of performance indicators and their use depends to a large extent on the model of management. PIs and rational scientific management are closely related, and even have something of a symbiotic relationship. Scientific management depends heavily on PIs, whereas the 'behaviourist' approach tends to rejects such 'simple' indicators of performance, and will instead rely more on subjective judgements.

The current predominately rational / scientific approach to educational management uses PIs to provide direction and control of the system, and although there are potentially many problems with this approach, there are also possible benefits. PIs can provide objective evidence to support, reject or quantify the various claims which are made about educational policy and practice. And this would seem preferable to a system which is based on various hunches and ideologies; in particular when set in the adversarial British political context. The next chapter explores some of these issues by looking from a historical perspective at the development of educational management and PIs

Chapter 5

The development of public sector and educational management

In the long run, this [National Literacy and Numeracy Strategy] may be the greatest dangers to establishing school cultures that can sustain growth and innovate problem solving far into the future (Earl et al)

Introduction

This chapter looks at the developments in public sector and educational management, which have led to the current testing and reporting regime; including KPIs and league tables. Firstly, the background to the 'need' for a new kind of management in the public sector is considered. This is follow by a description of the 'new public management', and the supporting legislative changes, in particular the 1988 education reform act and the Citizen's charter. All of these issues combined to raise the importance and significance of testing and assessment, in particular the results from key stage tests. The final section of this chapter details the main indicators which are made public in the official performance data in England and Scotland.

The 'need' for new management

Post-war Britain enjoyed a high rate of economic growth, which was largely due to receptive markets around the world and relatively cheap resources. This and the related prevailing mood of optimism was reflected in the education system. Resources were available to support the expansion and implementation of new policies and all of this occurred in what Bogdanor (1979) described as a climate of co-operation:

> Many policy decisions in education were taken over lunch at the National Liberal Club by a troika consisting of Sir William Alexander, Secretary of the Association of Education Committees, Sir Ronald Gould, the General secretary of the National union of Teachers, and the Permanent secretary of the Department of Education....such at least was the general belief.....
>
> (p. 160)

Docking (1996) makes a similar point with reference to Harold Dent's work;

> As recently as 1977, H.C. Dent could write that consultation and negotiation were the regular means employed in educational policy-making. What really made the English educational system 'tick', he said, was the fact that the various bodies who have to work in it - central and local administration, teachers and voluntary bodies - regard and treat each other as partners. He concludes by suggesting that if ever this 'all-pervading partnership' were to end, 'an entirely different - and much less happy - system would emerge.
>
> (p. xv)

However, changes during the 1970's in particular, forced a re-think of many of the perceptions and long held beliefs with respect to Britain and its place in the world. The new economies of Japan and the far east were fast becoming significant forces in the global economy, resources such as oil were becoming more expensive, and in Britain there was increasing industrial unrest. As pointed out in the last chapter, industry had to adopt new and more efficient methods of working, and new models of management. This put the 'spot light' on the public sector, which as Ranson and Stewart

(1994) point out was increasingly being viewed as inflexible and inefficient, and was subject to wide ranging criticism;

> Hospitals, and surgeries, schools and welfare agencies, local authorities and departments of government provide an array of services, benefits, subsidies and regulations that have formed the context of most people's lives. Nevertheless, for over a decade the public sector has looked beleaguered: its rationale doubted, its effectiveness and efficiency derided and its resources depleted...Public organisations have, purportedly, been self-interested and unresponsive to the public and have failed to fulfil the expectations of the post-war years of leading the creation of a fairer, more equal society.
>
> (p. 4)

Such views helped pave the way for a plethora of government sponsored reports which sought to make the public sector and local government more efficient and responsive to the needs of the community. In 1968 the Fulton Report on the civil service proposed that individuals and operating units should be held accountable for their performance by the use of objective indicators (Prince 1983). The Bains Committee (1972), in their report, *The New Local Authorities: Management and Structure*, made many recommendations as to how local authorities should be re-organised and which services they should provide. For example, it was recommended that they should adopt a corporate approach to enable the authority to provide a range of integrated services to meet the changing needs of the local community. As Pollitt (1993) points out the 'corporate approach' was not based on the current thinking in the British public sector, but rather had it origins in the business thinking of the US, and was applied to large multinational corporations. Lawton and Rose (1991) illustrate this in practice with the example of the creation of Chief Executive posts, which had the purpose of co-ordinating or 'joining up' the work of local councils. Similar cross-functional approaches have been adopted by elected members and their committees. Indeed the more recent (in the last few years) example of 'cabinet' structure within local authorities is a further continuation and development of this thinking.

More specifically in education similar arguments about the effectiveness of schools, and questions about the accountability of the teachers were being asked (eg Weinstock 1976). Callaghan's *Ruskin College speech* (Callaghan 2001) brought to the fore, the tricky question of 'who runs our schools', and in many respects marked the beginning of a *new educational management* (Baker 1994). A number of high profile issues such as, *William Tyndale*[69], set the scene for the then Labour Prime Minister to publicly question the performance and effectiveness of the education system (Batteson 1997, Demaine 1999). Although by present day standards the questions raised by Callaghan appear somewhat benign[70], they nevertheless struck a chord in many minds. Baker (1994) highlights the significance of the speech:

---

[69] A primary school in London which came to national prominence over its teaching methods.
[70] Certainly it contrasts in terms of style with Blair's Ruskin speech thirty years later (Blair 1996)

> It was a clever speech which marked an important turning-point. The need for schools to be accountable for the public money spent on them was now grudgingly accepted and the debate focused more on teaching methods and curriculum choice than on resources. He was asking the tough questions which too many educationalists had pushed to the back of their minds. He also avoided the pitfall of offering easy solutions. Critics of the many changes which have since occurred in education, notably under Mrs. Thatcher, have blamed Callaghan for unleashing the demons of centralised control and an over-prescriptive curriculum on schools.
>
> (p. 30)

However, whilst Ruskin laid much of the foundation for (or underlined) future Conservative and New Labour thinking, it was not until the third term of the Conservative government in the 1980s, with Kenneth Baker as the education minister, that substantial structural changes were made to the management and organisation of education (Chitty 1998). It was the arrival of Mrs Thatcher's conservative government which marked the beginning of a process which would begin to transform the public sector;

> The Conservative Government came to power in 1979 with a clear mandate for change: public services were to be made economical, efficient, effective and, most of all, accountable. Resources were to be managed instead of simply being used. 'Producer' power, in which the public service bureaucracies and professionals allocated resources as they saw fit, gave way to 'consumer' rights where customer sovereignty was to hold sway. This 'sovereignty' in the market place demanded that customers should be more informed about the services on offer so that they could make rational choices....Money supply was to be tightly controlled, tax cuts were to be made, and, in a much repeated remark, the 'frontiers of the state were to be rolled back'.
>
> (Johnson 1996: 187)

Many new initiatives were put in place to achieve this. Writers such as Pollitt (1993), Bourn (1994) and Broadbent *et al* (1996) highlight the importance of the Financial Management Initiative (FMI) which was launched in 1982. The white paper, *Efficiency and Effectiveness in the Civil Service* (Cabinet Office 1999b), aimed to;

> Promote in each department an organisation and system in which managers at all levels have:
>
> 1. a clear view of their objectives, and means to assess, and wherever possible measure, outputs or performance in relation to those objectives;
>
> 2. well defined responsibility for making the best use of their resources, including a critical scrutiny of output and value for money; and
>
> 3. the information (particularly about costs), training and the access to expert advice which they need to exercise their responsibilities effectively.
>
> (para. 13)

A New Public Management

Initiatives such as FMI and a 'new' philosophy towards the public sector, not just in the UK but in many other countries as well, contributed to a model of management which has been identified as *New Public Management* (NPM). This term was first coined by Hood (1991), although as he points out;

> NPM, like most administrative labels is a loose term. Its usefulness lies in its convenience as a shorthand name for the set of broadly similar administrative doctrines which dominated the bureaucratic reform agenda in many of the OECD group of countries from the late 1970s.
>
> (p. 3)

Although the term NPM is used in this thesis other writers have used different terms to describes what is essentially the same phenomena. For example, Pollitt (1993), refers *'Managerialism'*, and Lan and Rosenbloom (1992) *'Market-based public administration'*. More specifically in education, Chitty (1989) refers to the *'The New right and a New education system'*; Ball (1990) *'New managerialism'*, Grace (1995) *'New Right ideologies'*, and Power *et al* (1997) *'New Education Management'*. To help put NPM into perspective, Rhodes (1991) in his introduction to Hood's article illustrates some of the key features;

> The 'new public management' has the following central doctrines; a focus on management, not policy, and on performance appraisal and efficiency; the disaggregation of public bureaucracies into agencies which deal with each other on a user-pay basis; the use of quasi-markets and contracting out to foster competitions; cost-cutting; and a style of management which emphasises, amongst other things, output targets, limited-term contracts, monetary incentives and freedom to manage.
>
> (Rhodes 1991: 1)

For Hood (1991) the key driving force behind NPM was four 'magatrends' which had been evident in public management over the last fifteen years[71];

- A slow down or reverse in growth of public expenditure
- An increase in the privatisation of public services
- Increased use of information technology for public services
- Greater internationalisation of public management

(Based on Hood 1991: 3)

Many examples can be cited to support the last three of these observations. For example, the Public Finance Initiative (PFI) and Public Private Partnerships (PPP), on-line government information sources (such as Ofsted reports), and foreign companies being able to 'own' and control public utilities. However, there are difficulties in making a simple assessment[72] of the first point; possible examples might include increased individual responsibility (and cost) for dental care. The true

---

[71] ie 1976 – 1991

[72] For example should the % or GDP or National Income be used or the absolute figures themselves.

picture is somewhat confused by the government's 'creative' accounting techniques, such as 'off-balance sheet' funding to 'pay' for new schools (ie. PFI / PPP). Not in the same league as Enron, but as Railtrack has demonstrated the costs and more importantly liabilities remain with the state. Accounting techniques aside, public expenditure has in reality increased, in part due to the availability of technology and the needs of society. For example being able to carry out new medical procedures, or the expansion of higher education.

Hood defined seven doctrinal components of NPM, which he refers to as 'overlapping precepts', which are typical of public sector policy in the UK, Australia, New Zealand and other OECD countries. However, they are not all necessarily always present, and their proportion will vary between different countries and situations:

Table 5: Hood's doctrinal components of new public management

	DOCTRINE	MEANING	TYPICAL JUSTIFICATION
1	*'Hands-on professional management'* in the public sector	Active, visible, discretionary control of organizations from named persons at the top, 'free to manage'	Accountability requires clear assignment of responsibility for action not diffusion of power
2	*Explicit standards and measures of performance*	Definition of goals, targets, indicators of success, preferably expressed in quantitative terms, especially for professional services	Accountability requires, clear statement of goals efficiency requires 'hard look' at objectives
3	Greater emphasis on *output controls*	Resource allocation and rewards linked to measured performance; break-up of centralized bureaucracy-wide personnel management	Need to stress *results* rather than *procedures*
4	Shift to *disaggregation* of units in the public sector	Break up of formerly 'monolithic' units, unbundling of U-form management systems into corporatized units around products, operating on decentralized 'one-line' budgets and dealing with one another on an 'arms-length' basis	Need to create 'manageable' units, separate *provision* and *production* interests, gain efficiency advantages of use of contract or franchise arrangements *inside* as well as outside the public sector
5	Shift to greater *competition* in public sector	Move to term contracts and public tendering procedures	*Rivalry* as the key to lower costs and better standards
6	*Stress on private sector styles of management practice*	Move away from military-style 'public service ethic', greater flexibility in hiring and rewards; greater use of PR techniques	Need to use 'proven' private sector management tools in the public sector
7	Stress on *greater discipline* and *parsimony* in resource use	Cutting direct costs, raising labour discipline, resisting union demands, limiting 'compliance costs' to business	Need to check resource demands of public sector and 'do more with less'

Hood (1991, 4)

The influence of the hybrid schools of management is evident in many of these components. For example, Peter and Waterman's (1982) bias for action is similar to the first point of 'hands on' management and freedom to manage. The second point is closely related to the principles of management by objectives, and the emphasis on output controls and the need to stress results rather than procedures, is essentially the same as Hammer and Champy's (1993) argument of needing to organise around outcomes, not tasks. The importance of competition and the market (point five) is also stressed by Freeman (1993) with respect to TQM. Although the learning organisation approach and the Deming 'flavour' of TQM are significantly different to Hood's model, there are nevertheless some parallels; for example rewarding flexibility and breaking down barriers and monolithic structures. Indeed, Hood makes specific reference to 'proven private sector management techniques' (whatever they are?), implying that 'old public management' is based on the old bureaucratic models which have been largely rejected by business.

The NPM model fits well with many of the principles and aspirations of the new educational management. With respect to the first point, heads as the managers are clearly visible and 'named', and have a significant degree of managerial freedom; although as previously argued, far more so in a day to day, rather than a strategic capacity. Within schools quantitative goals and targets are very important, and are used to hold people accountable. And in the system as a whole great emphasis is place on quantitative goals and targets such as KPIs, which are linked to the allocation of the resources. The side-lining of the LEAs is an example of the break up of 'monolithic' units, as are the contractual type of relationships which exist throughout the system. A belief in 'the market' is evident in much of the government's underlying (and overlying) philosophy. Greater use is being made of flexible[73] employment principles, although it would not necessarily be correct to assume that these approaches have come from the private sector (see Morgan and Allington 2002). Furthermore, 'private sector' PR techniques are widely used by the government when communicating with 'their' customers (Doe 1999). Finally, the broad thrust of the educational reforms over this period has been aimed at raising standards (or rather those standards which are measured) without a corresponding rise in costs.

There are however strong arguments about the effectiveness and appropriateness of some aspects of NPM as a management model for the public sector[74]. Broadbent *et al* (1996) explore the issue of *principal – agent* theory (PA) as a method of public sector management. PA is closely related to or

---

[73] Two distinct interpretations can be made from the term flexible. For 'enlightened' employers it means individual employees being able to take many different roles in the organisation. However the government's view is much more 'hire and fire' with impunity; with for example throughout the education system the widespread use of agency workers and temporary contracts.

[74] It is of course the case that NPM is what Hood observed rather than what he advocated

indeed could be considered a part of NPM[75]. In essence one party (the principle) 'hires' another (the agent) to carry out particular tasks. This leads away from a partnership model, to one based on accountability and contracts; such a change in relationships is evident in the new educational management. LEAs, schools and teachers as the agents, are effectively contracted (and by definition held accountable) by the government (the principle) to provide 'educational services'. However, as Broadbent *et al* (1996) point out there are inherent problems with such relationships, particularly in terms of the different values of the various parties, and of the principle (the government) suffering from an 'information asymmetry'; ie. not knowing what the agents are doing. Indeed, as they point out this approach leads to the use of bonus systems and performance related pay to try and correct these problems, and the difficulties of using such systems in complex organisations is widely recognised (eg. Richardson 1999):

> In both the economics literature.... and the accounting extensions.... the solution to a PA problem involves the development of an optimal incentive package that furthers the objectives of the principal and constrains the actions of agents.
>
> (Broadbent *et al* 1996: 266)

This they point out leads to ever greater control and the development of punitive accountability systems, which reduce the level of trust between the principle and the agent, as well as between one agent and another (ie. between individual teachers or schools). They conclude that PA models (and by implication much of NPM) are not conducive to the effective management of much of the public sector;

> We question the use of this PA approach in the context of the 'caring professions' of education and medicine as a way to control professional activities... [and] we raise serious questions about the ethical issues of applying PA theory and the extent to which it is appropriate to implement 'economic reason' in the context of the public sector.
>
> (p. 259)

The Education Reform Act 1988

A significant amount of new legislation accompanied and supported the NPM trend, for example legislation had to be enacted to facilitate PFIs and competition in the public sector. In education too substantial changes were made, the most significant was the 1988 education reform act (ERA), which bought about many features of the present day education system. An important aim of this act, and the others which surrounded it, was to increase competition and accountability.

Early, in the new conservative parliament, the 1980 education act 'set the ball rolling' by allowing parents to state a preference in terms of which school they wanted their children to attend. To help

---

[75] Hood identifies PA as one of the ideas which contributed to NPM.

achieve this independent local appeals committees for admissions were established. The act also required LEAs and governors to provide information about schools, and that parents were represented on school governing bodies. The 1986 education act made changes to the composition of school governing bodies, in particular increasing the number of parent governors, at the expense of local authority appointed governors.

However, it was the 1988 act which was to have the greatest impact on the 'modern' education system, and in particular 'who' was in control. Basini (1996) points to the seemingly contradictory message in the act of increasing central control whilst at the same time increasing local control. Chitty (1989) too, traces the developments leading up to the act makes a similar observation:

> The central purpose of the Education Reform Act is that power should be gathered to the centre and, at the same time, devolved on to school and parent, both processes being at the expense of mediating bureaucracies, whether elected or not.
>
> (p. 219)

Significantly, these changes severely reduced the involvement of the LEAs in the management and organisation of the education system, in effect from being 'partners' to something more akin to agents[76]. As Baker (1994) observes, the government now had 'new' partners; however they were not from within the systems;

> The scope of the Bill was enormous, amounting to a major reshaping of the education system. What made it so radical was that change was initiated and imposed from outside. Not only was the driving ideology derived from the think tanks of the so-called New Right but even the ministers themselves thought and spoke as if they were 'outsiders', rather than as partners within the education system. Even the Department of Education was regarded as untrustworthy by the small group of people who were drawing up the changes.....This was a piece of legislation designed and implemented by 'outsiders' intent on breaking down the fortress walls of education.
>
> (p. 3)

Cave (1990) points out that the act was designed to break down the possible (probable?) resistance from within teaching world due to notions of 'professional autonomy'. This would be achieved by forcing a market type approach, as well as reducing the opportunity for maintaining and developing alliances between schools and LEAs;

> The Act [1988] is arguably the most radical and certainly the most comprehensive attempt ever to redesign the education system..... The means by which standards are to be raised are clear: it is to be achieved by restricting professional autonomy, by increasing parental power, by exposing the system to the forces of open market competition and by streamlining the governance of schools through curtailing the role of the local authority middle tier and greatly increasing the powers vested in the secretary of State for Education.
>
> (p. 99)

---

[76] As in P-A theory

Three areas from the ERA are particularly pertinent to this thesis; the nationalised curriculum, choice or pseudo - market principles, and local management. Details of what is contained in the national curriculum are considered in more detail later in this chapter, however the main principle was to ensure that all children received basically the same broad and balanced curriculum. However, as Basini (1996) points out, due to the speed of its introduction it has had to be substantially revised[77], in part due to inherent technical weaknesses, as well as to shortcomings in the implementation process. The act further developed the principle of 'open enrolment', with schools having to accept a child if they had the space: assuming the child had the appropriate aptitude. This created a sort of market principle, for *some* parents and their children. Others for example which could not afford the necessary transport costs were effectively excluded from this market. Significant changes were also made to the funding arrangements of schools. The basic principle, which has continued to the present day, is that schools manage their own resources; ie. local management of schools (LMS). Prior to 1988 the LEAs had been closely involved in these activities.

Significantly, the act was implemented very swiftly and with virtually no consultation, and as Baker (1994) points out from his subsequent discussions with the architect (and his namesake) this was entirely intentional and designed to 'wrong foot' the opposition:

> The 1988 Education Reform Act knocked schools, teachers and parents into a spin. Most are still rotating or, at best, still dizzy from the effect. It was a revolutionary Act brought in and then implemented at revolutionary speed. There was strong opposition to it from organisations representing parents, teachers, the Churches, the local authorities and many other groups. Yet the pace and extent of change were so great that they had difficulty mobilising and fighting across so many fronts. They were quickly overrun. Only after most of the reforms were in place did the opponents manage to regroup and present a united appearance. Then as rebellion broke out, most notably over the new school tests, it was the government's turn to reel and the whole edifice of the school reforms looked shaky.
>
> Baker (1994: 33)

Before considering the assessment and testing aspects of the ERA, the citizen's charter is briefly discussed. As well as supporting the general thrust of the government's policies, the charter made specific demands in terms of testing and reporting performance of schools and many other areas of public service.

The Citizen's Charter

The conservative government, under a new leader - John Major, produced the white paper, *The Citizens Charter: Raising the Standard* (Cabinet Office 1991), which launched the 'Citizen's

---

[77] Eg. Dearing review in 1993

Charter'. This was aimed at both improving public services, and importantly from the perspective of this thesis, providing the public with information about the performance of the different services, such as health education and transport. Leach (1996) states that: "the role of the Charter was essentially to publicize those policies (eg. choice of schools) and further assist their implementation." (p. 107). Kouzim *et al* (1999) go further and state that: "The ultimate purpose (of the citizen's charter) is to renew citizen trust not only in public services but also in the state." (p. 127). And in a world of 'spin doctoring', Kingdom (1996) asserts that: "It offers a government embarrassed by shortcomings in state services another advantage: scapegoating". (p. 20)

Reference is made here to both the original Citizens Charter (1991) and the subsequent *Our Children's Education: The updated Parents Charter* (DES 1994). In terms of education these documents are very similar, with the latter reflecting the legislative changes which occurred in between 1991 and 1994. In Scotland The *Parent's Charter* (Cabinet Office 1995) is very similar to the English (and Welsh) version. The Citizen's Charter has four main aims:

> **Quality** - A sustained new programme for improving the quality of public services
> **Choice** - Choice wherever possible between competing providers is the best spur to quality improvement.
> **Standards** - The citizen must be told what the service standards are and to be able to act where service is unacceptable.
> **Value** - The citizen is also a tax-payer; public services must give value for money within a tax bill the nation can afford
>
> (Cabinet Office 1991: 4)

Although the term 'improving quality' is used extensively throughout the documents, and indeed in many government publications and proclamations; as Elcock (1996) points out it is a rather elusive concept, which is not really defined anywhere in the charter. There is an overall impression of a general or 'holistic' rise in standards, but in practice it appears to be far more narrowly defined: ie. as exam or test results;

> Better information about schools is also important for raising standards. For example, publishing tables which compare the performance of schools has encouraged many schools to take a hard look at the examination results their pupils achieve and how the school can help the pupils to do better.
>
> (DFE 1994: 3)

The concepts of markets and choice is central to the charter and much of the government's philosophy and ensuing policy. For example, the White paper *Competing for quality*, which was published by the treasury under the auspices of the Citizen's Charter (Leach 1996), states that, "competition is the best guarantee of quality and value for money" (Treasury 1991, p. 1). It would however be wrong to assume that the charter introduced the competition principle. As Harrison (1995a) points out the legal requirement for choice and competition had already been established in 1988, with the compulsory competitive tendering (CCT) process which was aimed at local

authorities. Furthermore, Chandler (1996) points out the term and many of the principles of 'charters' can be traced back to previous schemes established by local authorities. in particular York City and Lothian Regional Council[78].

The 'charter' approach is consistent with many aspects of the hybrid management, for example claims such as; "The Government is now determined to *drive*[79] reforms further into the core of public services" (Cabinet Office 1991: p. 4). is very much in harmony with BPR. It also chimes with Hood's definition of NPM, which highlighted the use of rivalry (or competition) as the key to better services, as well as the importance of PR to NPM.

In 1994, *The updated Parents Charter* (DES 1994) was very widely advertised, and distributed to all 20 million homes in England and Wales (with and without children). It gave details of what the government saw as its achievements in education, and an indication of future plans. Most relevant to this thesis, was the inclusion of future targets and details of the documents which would be published or made available, these included;

- An annual report for each child, which includes national examination and test results, attendance and attainment in non-tested subjects and other activities.
- Reports from Ofsted inspections.
- Annual performance tables of examinations and national tests (league tables)
- School prospectus
- Annual report by the school governors

Although the charter (or charters) were very high profile publications, their real impact is questionable. For example in education, Leach (1996) points to evidence that suggests a low level of awareness, including the government's own research which found that only 29% of parents were able to recognise the booklet eight weeks after distribution. With respect to local government, Harrison (1995b) found that 82% reported that the citizens charter had had "no effect" on their work. This finding may be partly explained by the fact that many local authorities had already implemented many of the relevant principles before the charter itself was published. However, she does point out that 59% of the authorities found that the publication of performance indicators had had 'some effect' on their work and policies. Furthermore, Kouzim *et al* (1999) who looked at 'charters' in a number of different countries, question whether the 'charter' concept is fundamentally flawed;

> Charters sound very convincing at first sight but there are two in-built contradictions in the basic concept. The first inconsistency of citizens' charter is that it makes allusion to the *English Magna*

---

[78] Although these were aimed more at encouraging and enabling public consultations
[79] Emphasis added

> *Charta Liberatatum* of 1215, which was a letter of privileges for noblemen rather than a catalogue of liberal rights for all citizens... When charters are used in a new ideological context, such as with "new public management", the meaning of charters changes significantly. It is no longer a catalogue of rights and duties of the ruler and the ruled, but a "quality" checklist for clients of public services. This is especially true in the British Citizens' Charter which builds on the user service provider relationship instead of citizen-state relationships.
>
> Second, charters are based very much on older social contract theories (Rousseau, Hobbes, Locke) according to which citizens believe, unconditionally, in the legitimacy of the state as long as the state provides security. This is no longer true today where the market contract becomes more dominant than the classical social contract.
>
> (p. 128)

The underlying principles of charters continue to the present day in a number of forms. The same sorts of promises are made in the numerous policy manifestoes and pledges, in many cases with specific targets, or the more general and increasingly ubiquitous 'zero-tolerance' assurances. In terms of providing information, Power and Clark (2000) illustrate how the charter principles have continued to be applied, and have for example led to new labour's white paper, *Excellence in Schools* (DfEE 1997). This introduced Home – School agreements; indeed this 'charter principle' might also be viewed as 'conferring privileges on the noblemen', rather than the citizens.

The rise of Assessment and Testing

Murphy and Wilmut (1999) point out that the key shift in policy in terms of assessment and testing has been from being teacher led to government imposed;

> Until the second half of the 1970s educational assessment attracted relatively little attention. There was a general consensus that teachers were responsible for teaching and assessment in the classroom, and that examining boards provided external certification. In contrast, assessment practices and outcomes are now matters of public debate.
>
> (p. 116)

This statement is true with respect to the relatively recent history of education, however in the more distant past testing and assessment have been an important part in the management and organisation of the education system; moreover it would appear that much can be learnt from these earlier experiences. The 19th century 'Revised Code' marks an important starting point in the principle of high stakes testing, and even shares some of the underlying principles of modern KPIs.

The impetus for the code came from concerns that Britain in the mid nineteenth century was becoming less competitive in the world, and that this was largely due to the poor standards of education for many children. A Royal commission was set up, and in 1861 the *Newcastle Report*'s main proposal was that schools should only receive a grant if they have been given a satisfactory report by HMI. The main criteria was the satisfactory performance of the pupils in reading writing and arithmetic tests. This system of 'payment by results' has been widely criticised for causing a

narrowing of the curriculum, or confining it to what is tested, as well as encouraging teachers to 'teach to the test'. In 1867 Matthew Arnold observed that as a result of the Revised Code:

> The mode of teaching in the primary schools has certainly fallen off in intelligence, spirit and inventiveness. . . In a country where everyone is prone to rely too much on mechanical processes and too little on intelligence, a change in the Education Department's regulations, which, by making two-thirds of the Government grant depend upon a mechanical examination, inevitably gives a mechanical turn to the school teaching. . . and must be trying to the intellectual life of a school.
>
> (Cited in Baker 1994: 76)

This system of funding continued until the 1902 Education Act which introduced the block grant system for schools[80]. This was quite similar to the present system of formula funding, with the amount of money a school receives being largely determined by its size. However, the current trend is back towards some degree of payment by results or performance related pay, and as such is entirely consistent with the principles and theories of NPM and scientific management. At the individual teacher level the PRP system currently takes the form of 'Threshold' payments (DfES 2001c), and for schools the 'Schools Achievement Awards' (DfES 2001c). Both of these systems rely heavily on KPIs to judge levels of performance. Indeed in the case of school achievement awards one of the criterion is how far 'up the league table' the school has progressed[81].

In the early part of the last century concerns were voiced about the overly narrow secondary school curriculum. This was heavily influenced by the 'grammar school' system, which it was argued did not meet the needs of industry and commerce (Baker 1994). This issue was addressed by the Spens Report in 1938 ,which advocated a tripartite school system (Grammar, Technical and Secondary modern). The eleven plus was designed to make an objective assessment of children's innate intelligence, and from this the child would be sent to the most appropriate school. Concerns over the accuracy and reliability of the tests, as well as a general liberalisation of the education system during the 1970's led to the decline of the selective system, and the development of comprehensive schooling, and therefore in many primary schools the 11+ has become redundant.

Although some selective schools still exist and pupils wanting to attend have to sit the 11+, there are however a growing number of comprehensive schools which are in all but name selective, and for many of these a far more insidious form of selection is used. For example, selection may in reality be based on post codes, and more subjective assessments which claim to measure aptitude (Webster and Parsons 1999, Hattersley 1998). Indeed, as Mooney (1999) warns: "The 11+

---

[80] Significantly this act also established LEAs which were to have a major influence on the 'local' curriculum (Phillips 2000)

[81] There are more subjective criteria for small and special schools whose performance is not reported in the performance tables

examination was abolished in most areas many years ago. We are now in danger of seeing it reappear in a grossly mutated form with little control" (p. 1).

With the demise of the revised code at the beginning of the 20th century, Baker (1994) points out that the responsibility for the curriculum and standards was largely in the hands of 'the teachers'[82], and it was not until 1974 with the setting up of the Assessment of Performance Unit (APU) that the government would again begin to gain 'control'. Pring (2000) states that the APU; "aimed to provide a curriculum framework and a longitudinal comparison through which standards in schools might be measured over time." (p 4). And as Shaw (1996) points out, the unit's purpose was to; "*monitor* pupil performance and to *inform* the Education Secretary of educational standards and any discernable trends in those standards". In effect the APU had a monitoring and reporting function rather than a controlling function. Baker (1994) argues that this low key approach, which had little significant effect on the curriculum, contributed to its downfall; "it [APU] was perceived as a failure by the Conservative government since it gave ministers no control over the curriculum"(p 81). Indeed were it still in existence today it could in theory help resolve the conflicting and highly politicised debate as to whether standards have risen, or simply made to *appear* that they have (Bassey 2001).

Whilst the underlying philosophy which led to national testing programmes can be seen in the development of political thinking as well NPM since the 1970's, Shaw (1996) suggests that the first clue of what was to come, and what would form the basis of the present day system, was a speech by Keith Joseph, the then Secretary of State for education:

> At the North of England Education Conference in Sheffield in 1984, he [Joseph] emphasised:
>
> - a commitment to higher standards
> - examinations that focused on pupils' achievements (rather than on simply selecting or ranking them)
> - the inclusion of parents and employers in the selection of educational objectives.
>
> (p. 30)

At face value these are of course very reasonable and quite appropriate aspirations for schools. However, in practice they did pave the way for more explicit and radical solutions to the 'problems' of the education system.

Task Group on Assessment and Testing (TGAT)

Three years later, Kenneth Baker the then Education secretary, established the *Task Group on Assessment and Testing* (TGAT), which was to advise on the assessment arrangements for the National Curriculum. Much of the current testing and assessment systems, and therefore KPIs is

---

[82] ie the educational establishment

based on TGAT's work. The group established four main criteria which would guide and inform their recommendations:

> - The assessment results should give direct information about pupils' achievement in relation to objectives: they should be criterion-referenced.
> - The results should provide a basis for decisions about pupils' further learning needs: they should be formative.
> - The scales or grades should be capable of comparison across classes and schools, if teachers, pupils and parents are to share a common language and common standards: so the assessments should be calibrated or moderated.
> - The ways in which criteria and scales are set up and used should relate to expected routes of educational development, giving some continuity to a pupil's assessment at different ages: the assessment should relate to progression.
>
> (DES 1987 para. 5. Cited in Shaw 1996)

Central to their report in 1988 was the importance of using the results for formative assessment. However, this was to set about a clash of philosophies, with on the one hand educationalists wanting such information to improve their teaching, and on the other, the government with its NPM approach of using accountability and market forces to 'drive up' standards (ie summative testing).

As Stobart (2001) points out the report proposed that National Curriculum assessments should be based exclusively on Teacher Assessments, which would be externally moderated; in effect formative assessments[83]. The report was welcomed by both the teaching unions and the labour party, but not Mrs Thatcher, the Prime Minister (Baker 1994). The original TGAT proposals have been substantially modified over the years (Daugherty 1995), not only by the then conservative government, but also the new labour government (Webster and Parsons 1999). Now as Stobart (2001) points out the key purpose of National Curriculum assessment is to hold teachers and schools accountable, with very little if any notion of formative testing[84]. The results do of course play an important part in the planning process, for example with the use of the *Autumn Package* (DfEE 2000b). Furthermore, Ashby and Sainsbury (2001) looked at how schools used the returned exam papers for planning purposes, and found that they were used for curriculum planning, although more use was made of the English and maths tests rather then science; this they attribute to the greater pressure to meet literacy and numeracy targets. The important question still remains as to whether feedback from NC assessments is used for formative assessments in the true meaning, (ie. to improve the education for individual children) or as a means of helping schools meet 'their' targets.

---

[83] Summative assessments were still to used at 16, ie. for external examinations

[84] He cites the Rose Enquiry which makes no mention of any formative role for NC testing

The debate over using assessment for both formative and summative purposes has been going on for many years (Gipps 1994), and it is perhaps a little naive to believe that assessment can simply be used for both purposes. As Stobart (2001) points out the summative function will tend to dominate;

> ...there is general agreement that, where there are both formative and summative purposes, there will invariably be a shift towards more emphasis on those summative functions which inform *managerial* concerns for accountability and evaluation
>
> (p. 31)

Similar observations have been made in other areas of management. For example, there is much evidence that individual assessment systems can not be used effectively for developing staff (formative) and for remuneration (summative)[85]. Therefore it would seem unreasonable to expect equally valid formative and summative assessments from just one system (ie National Curriculum tests), indeed trying to do so may have a negative impact on the teaching and the school curriculum.

Furthermore, as Murphy and Wilmut (1999) point out educational assessment (or indeed any other form of public sector performance appraisal) is inextricably linked to the broader organisational and political context, and it is therefore important to consider this wider picture; ie, what is the purpose of the assessment and how does it contribute to the education process;

> Firstly, educational assessment is a complex activity which can never be undertaken without facing major challenges related to accuracy, dependability and impact on learning. Secondly, a natural consequence of this is that decisions about how educational assessment is to be undertaken are likely to be highly political. Powerful interests are invested in arguments for carrying out assessments in particular ways. For this reason assessment reforms need to be made hand in hand with curricular reforms, in relation to clearly defined ideals for lifelong learning.
>
> (p. 137)

The TGAT proposals and the many subsequent changes which established the new assessment arrangements were the cause of intense conflict, indeed the conflict is still evident today. Mike Baker, who at the time was the BBC political correspondent (prior to becoming education correspondent) was well placed to observe and interpret many of the issues;

> It is interesting to reflect that, in England and Wales, there has been far more fuss over testing than over the curriculum. While that may seem a curious priority, it also reflects the realisation that control over testing leads to control over what is taught. That is why the struggle over testing has been so fierce. The TGAT report, by putting the stress on teacher assessment and formative tests, was a serious blow to the government. It gave ownership to the teachers just as the government was trying to reduce that level of ownership. It meant the way forward was bound to be messy, as the government

---

[85] See for example Fletcher (1993) who summarises 'best practice' which advocated separating the 'money' from the 'development' in performance appraisal systems

> tried to wrench the national tests away from the increasingly complex teachers' model to a simpler, straightforward model which could be used to compare school with school. It set the stage for the row which reached its climax in the teachers' test boycott in the summer of 1993.
>
> (Baker 1994: 105)

Following the 1988 Act and the TGAT report, there were a number of changes to the education secretary, and therefore who would have task of 'driving' though the new testing system. Firstly, John MacGregor replaced Kenneth Baker who had implemented the 1988 Act. McGregor's stewardship marked a significant change in the relationship between the government and 'teachers', and yielded a number of concessions on the testing arrangement. However, he was replaced by Kenneth Clarke who was a very different 'kettle of fish'. Clarke 'delighted' in conflict (Baker 1994) and preferred to call a 'test a test' and not some 'fancy assessment'. Then in 1992 Clarke was replaced by John Patten, who had quite recently entered parliament (in 1979), having previously had an academic career, and ironically *should* perhaps have been an ally of the 'teachers'. In terms of assessments, Patten's first difficulties arose over marking and grading of the relatively new GCSE exams[86]. Importantly it was not only the teachers who were offended by his handling of the situation but parents as well. There was even disagreement between the government and the Schools Examination and Assessment Council (SEAC) (appointed by the government) over the content and purpose of testing and assessment.

It was however the planned introduction of KS3 tests for 14 year olds which was to prove the biggest battle and the downfall for Patten. In 1992 these were voluntary, but in 1993 it was decided to include English tests alongside Maths and Science. Furthermore, the results from of these now compulsory tests would be made public, from which league tables would be compiled. Moreover, these tests were to be 'short unseen written examinations', and significantly would take precedence over teacher assessments.

The inclusion of the English tests proved to be a critical factor in mobilising the teaching unions in a rare act of unanimity in opposing and boycotting the tests. This was in spite of a last minute offer not to publish the English results. To try and 'help' the government out of the mess, the London Borough of Wandsworth unsuccessfully sought an injunction against the boycott. It is ironic, or perhaps significant, that some years later in 1999 Wandsworth should find themselves 'in the dock', found guilty by a QCA investigation of giving 'their' schools stolen details of the KS2 English exams (Cassidy 1999). Battle continued during the teacher's conference season of 1993 and in spite of a promised review of the National Curriculum (Dearing Review), and even that the news media and opinion polls were supporting the 'teachers' case, the tests went ahead in June

---

[86] The substantial rise in performance was viewed by some as suspicious, furthermore they were viewed by some as ignoring the 'traditional' basics.

1993. However, they were something of a disaster and provided no meaningful data, with fewer than 150 schools administering the tests and completing their returns to SEAC.

Although this episode may represent something of a 'victory' to the educational establishment in terms of formative assessments over summative tests, 'league tables' for KS4 in the form of GCSE were still published in 1993. Likewise the KS2 results have been published since 1996. Indeed, the current government is 'preparing the ground' for the publication of KS3 results in 2002, by emphasising the substantial progress made up to KS2, for which tables are published, and the relative lack of progress between this and KS3, for which tables are not produced (Morris 2002).

The experience of testing and assessment, and the publication of the results in Scotland contrasts with the English (and Welsh) experience. In spite of a long history of 'English' control, Scottish education is different in many areas, and has always sought to maintain its independence (MacKenzie 1999). The government, with a staunch Thatcherite in the form of Michael Forsyth sought to introduce tests in 1991 which were similar to those proposed for England, and which would have been used to compile league tables. However, this was met with very strong opposition from the main teaching union (Educational Institute of Scotland which represents 80% of teachers), parents, academics and education authorities (Saunders 2000). Indeed, Munro and Kimber (1999) point out that 66% of parents nationally withdrew their children from the 1991 tests, and given the run up to a general election (which was expected to return a labour party opposed to this form of testing) a significantly different form of testing was agreed; the *5-14 National Assessments*.

A review of the curricula, and testing in England and Scotland

The details below refer to the system which existed in 1999 – 2000, the period when the research for this thesis was carried out. A number of alterations have occurred since, and where relevant will be discussed later in the thesis. Although some points made below have already been made they are repeated in this section to help maintain continuity. Furthermore, only the arrangements for mainstream state school children aged five to sixteen are considered, although in general the independent sector is adopting the same systems.

In England the National Curriculum which specifies what children must study, and the levels of performance they should achieve at different ages is compulsory in all state schools. It is set by the Qualifications and Curriculum Authority (QCA), which is appointed by the Government. The Office for Standards in Education (Ofsted), is responsible for ensuring schools adhere to the curriculum.

The national curriculum aims to raise overall standards of education by ensuring that all children receive a broad and balanced curriculum. In specific circumstances parts of the curriculum may be dis-applied from certain children. By specifying what is taught the curriculum makes it easier for children to change schools, and for teachers to be 'trained' and kept up to date. Furthermore, material and resources can be specified and supplied on a national basis, for example, key stage test revision resources.

The curriculum is divided in to four Key Stages. At the primary level KS1 is for children aged five to seven, and KS2 for those aged seven to eleven. Tests in Maths English and Science are taken on a set day in an examination style at the end of KS2, and are marked on a 1-6 scale (levels). Teacher assessments are also carried out for these three subjects as well as for other curriculum areas. In secondary schools KS3 is for children aged eleven to fourteen and KS4 for those aged fourteen to sixteen. For the final two years (KS4) students will typically follow exam courses leading to a number of GCSEs, and grades ranging from A* to G may be awarded. Various vocational courses may also be taken.

There are three core subjects, English, Maths and Science, in the national curriculum, which are compulsory for all children aged five to sixteen. In addition, there are foundation subjects which include; technology, geography, history, music, art and a foreign language. The foundation subjects may or may not have to be studied depending on the particular key stage, and the child's preferences. Furthermore, the National Literacy and Numeracy strategies (NLNS) for primary schools provide very specific instructions on what is taught, and how this should be carried out.

In Scotland the actual curriculum is not defined in law, ie. there is no 'national curriculum' as such. The precise content and management are the responsibility of the individual education authorities and the specific head teacher. However, in practice the bulk of the curriculum is centrally defined. The Scottish Consultative Council on the Curriculum (SCCC)[87] which is appointed by the government is the principle advisory body to the Secretary of State. As in England, schools and the curriculum they provide are inspected by HM inspectors of schools.

The actual curriculum can be considered in two parts, firstly for ages five to fourteen, and secondly from fourteen to sixteen. Children attend primary schools from the age of five to twelve, and secondary from twelve to sixteen. Therefore, the 5-14 curriculum overlaps these two phases. The subjects covered by the 5-14 curriculum are similar to the English national curriculum, in that it

---

[87] Now Learning and Teaching Scotland

aims to provide a broad balanced education. However, the national guidelines give more general details about the subjects rather than detailed instructions on the range of skills, knowledge and understanding which have to be taught. In the final year of primary school pupils (aged 12) do National 5-14 assessments in Maths Reading and Writing. This is on an individual non-examination basis and when the teacher feels the particular child is ready.

The 14-16 stage is quite similar to England, with the actual curriculum being largely determined by the exams (Standard Grades) the child is taking. These exams are broadly similar to GCSEs, in addition there are also various vocational options.

In both England and Scotland testing and assessments are used throughout the education process, from Baseline assessments for children entering the system to GCSE / Standard grade exams. Whilst much of this testing is used for 'internal' or formative purposes, some of the tests (KPIs) are used for target setting and are publicly reported in performance tables. All the official targets and results are based on whole school aggregates. The official publications in England give details of the individual schools and others in their LEA, with separate publications for each LEA. In Scotland the details for all schools are in one publication, which is arranged on an EA basis.

Below is a summary of the main official publications for the 1998 results (ie. current at time of research):

- 1998 Secondary School Performance Tables, for each LEA, published by DfEE
- 1998 Primary School Performance Tables Key Stage 2 Results, published by individual LEAs
- Examination Results in Scottish Schools 1996 -- 1998, Published by the Audit Unit, HM Inspectors of Schools
- Attendance and Absence in Scottish Schools 1995/96 to 1997/98, Published by the Audit Unit, HM Inspectors of Schools

From these official publication the media or any one else is free to re-publish the data in any format, including rankings or league tables. Below are details of the main indicators from these various publications:

For English secondary the main published data for each school includes details of the number of pupils and those with special educational needs. The number taking GCSE / GNVQ at the end of KS4. The percentage obtaining 5 or more grade A* - C, 5 or more A* - G, 1 or more A* - G, and the average point score. The improvement (change) over the past 4 years is shown for the number of 5+ A* - C and 1+ A* - G. Where appropriate, details are given of the percentage passing vocational qualifications. The percentage of half days missed due to authorised and unauthorised

absences is also given. Averages of these figures are also provided for the LEA as well as Nationally.

In English Primary schools the total number of pupils and those with special educational needs is given, and the number of those aged 11 and therefore taking KS2 assessments is also given. Where this exceeds 10, the percentage of those achieving level 4 or above is given for both the Teacher Assessment and the Tests in English Maths and Science. Averages of these figures are also provided on an LEA and National basis.

In Scottish secondary schools details are grouped by Education Authority (EA). The total number of pupils on roll is given. The percentage of S4s gaining 5+ standard grades at 1-2, 1-4, 1-6, is given for 1998 and the two preceding years. Details of the percentage gaining National Certificate modules and the percentage staying on post S4 are given. Details are also given for Higher grades and S5/6 results. Comparative EA and National data is also provided. In the attendance publication information is given for authorised and unauthorised absences, again this is on a school / EA basis, and includes historical data.

No data is published for Scottish primary schools, although the results from the National 5-14 assessments are passed on to the EA and HMI as well as individual parents.

In summary, at the secondary level very similar data is published for schools in England and Scotland. There are some similarities between the English primary and secondary schools, however at the primary level the two countries are very different with no school level data being made publicly available in Scotland.

Educational League tables

The governments current position regarding the publication of performance data, and the consequential re-publication of the data in league table format remains essentially the same; namely that they are very much in favour. For example, in England, Mike Tomlinson, the Chief inspector of schools (in 2001) states that:

> Parents today are better informed than ever about our schools and the quality and variety of provision within them.... Successive governments have been determined that parents should be equipped with the right tools to make critical choices about their children's education.... This move towards greater public scrutiny of our school system - partly driven by the media increased the flow of information to parents.
>
> (Tomlinson 2001: 1)

Significantly, he draws attention to the important role of the media, which is entirely consistent with the adoption of NPM and the use of PR. Officially the government only publish the data in alphabetical format, and once this is in the public domain, their responsibility ceases; a point made by Jacqui Smith as Education Under-Secretary

> I believe strongly that the tables produce information about the relative performance of schools, although I fear what local and national newspapers might choose to make of the information. The Government can take some responsibility for what is published. However, if we have the power to control the message that goes out about schools, please believe me-we would use it.
>
> (HOC 2000: Column 66WH)

This view is a little disingenuous and does not of course explain the rush, after the publication of the league tables, by ministers to be 'photographed' with the schools at the top of the league, or indeed the honours they like to bestow on the respective headteacher. Furthermore, the government itself uses 'crude' league tables to grant 'school achievement awards' (DfES 2001e). Sums of several thousands pounds are given "to the highest-ranking schools" (p. 3), based either on improvement or highest performance. In reality the ranking of the performance data in league table format, which is widely publicised, is very much what the English government wants, and what they believe will 'drive up' standards.

In Scotland too, Jack McConnell (Henderson 2001), then Education Minister[88], justifies his strong support for the publication of performance data, although as in England the government point out they do not publish league tables. Once again this part of the process is left to the (helpful) media'.

It would of course be wrong to assume that the media publish performance data or league tables just for the benefit of the government or for other altruistic reasons. Clearly, there are commercial interests, or put another way they are of interest to much of the public, and they therefore 'sell copy'. From the human perspective, Oswald (2001) suggests that;

> All over the world, school-teachers pin these tables up on notice-boards. Human beings are fascinated by rankings. It is presumably for Darwinian reasons: your many-times great-grandfather came high up a pecking order and that is why you are still here. Even in games, like football or table-tennis, people have to construct leagues.
>
> (p. 5)

Given the interest or potential interest and the benefit to sales, the media are very keen to publish the data. In England the necessary details are readily available, however a case in Ireland demonstrates the lengths the media will go to. The Sunday Times, and other newspaper fought a complex legal case to eventually secure access to secondary school examination results. (See OIC 1999). The case provides an interesting insight in to the arguments for and against publication, and

---

[88] Now First Minister

includes evidence from Chris Woodhead, the then Chief Inspector of Schools, supporting the media's arguments.

Although a high proportion of the media do publish school league tables, they are able to put their own slant on the data. The *BBC* website (www.bbc.co.uk) for example gives league tables, for all of the English primary and secondary schools as well as the Scottish secondary schools. As would be expected the results are presented in a 'matter of fact' way, and are essentially just a listing of the official data in a league format . English secondary schools are ranked on the percentage of pupils achieving 5 or more A* to C GCSEs, the primaries on the percentage of pupils gaining level 4 in the SATs exams, and for Scottish secondary schools, the percentage of pupils gaining 5 or more level 4 standard grades.

Other publications produce their rankings in other ways. For example, a local or regional paper such as the *Yorkshire Post*, gives ranking for the schools in their area based on the total point score. This they argue is fairer than the percentage of 5 A* to Cs, and significantly each school is given a few paragraphs to 'explain' their results, or put them in to context. On the other hand the *Sun* newspaper takes a somewhat different line. They give a number of mini leagues which show, for example; (the), Top 10 Comprehensives, Top 10 Grammars, Bottom 10 schools, Highest failure rate[89], worst 10 school for truancy. In many ways these contrasting approaches by the two papers demonstrates the difference between, a fairly genuine attempt to providing the public with information, and sensationalism to sell papers.

The tables published by the Sunday Times for independent schools provide an interesting insight in to the likely development. They publish a table, *The best value among the top 50 schools* (Sunday Times 2001). This gives the fees and number of A/B grade GCSEs and simply divides the two to give a 'Cost of A/B grades per year'. It is difficult to see how this might be of any real use, nevertheless, as will be discussed in chapter 14, similar proposals have been made for the state sector.

Conclusion

This chapter has outlined the background to the claimed need for a new kind of public management. Significantly, many of the arguments are similar to those in the previous chapter, and which led to new (hybrid) models of management. Indeed, 'old' public management corresponds closely to Weber's theories of bureaucracy, and both systems are considered by many, as

---

[89] % with no passes

inadequate to meet the demands of an increasingly competitive world. The answer to this apparent deficiency has been found in new public management, with its emphasis on performance measurement and accountability. Within education this change in philosophy has seen teachers move from being trusted professionals with a high degree of autonomy, towards something more akin to highly controlled and organised 'curriculum technicians'.

The development of performance indicator systems and KPIs has been accompanied by a degree of confusion as to their main role. On the one hand 'educationalists' want to use then for formative assessments to improve their teaching, and on the other, the government want to use them to control and hold individuals accountable, in the belief that this will 'drive up' standards. Ironically, both sides of the argument want essentially the same thing (self interests on both sides excepted): namely a better education system.

There can be little doubt that the indicators themselves, and just as importantly the broader management contexts within which they exist, will have significant behavioural consequences, which will affect the quality and effectiveness of schools. From this perspective the next chapter looks at some of the evidence and arguments about the intended and unintended consequences of 'high stakes' indicator systems.

Chapter 6

Performance Indicators and business accounting theory

Nothing could be worse. The evil effect of the Baldridge guidelines on American Business can never be measured (Deming)

(Baldridge National quality award guidelines look rather similar to the School achievement award guidelines)

Introduction

Chapter 2 outlined some of the similarities between the business and education world, in terms of measuring performance, with both facing significant challenges in being able to paint a true and fair picture. The issue of the two sectors using 'accounts' for internal development and external reporting was mentioned, as were the difficulties of developing reliable and valid measures, which compare 'like with like'. There is also a need to use a common currency throughout the respective sectors, (eg. exam passes or financial), as well as similar difficulties in identifying 'cause and effect', for example who made 'that sale', or who taught that child graphs. There is increasing convergence between the two sectors due to the increasing private sector involvement, and this will increasingly blur the distinction between the two accounting systems. For example, some companies (eg. RM computers[90]) profits are to a degree tied to academic performance measures.

As with education, businesses have many indicators that can be used to measure performance, although by definition *Profit* is still generally accepted as being the most important. As previously argued this is similar to a simple definition of value added[91], in that it is describes the difference between a starting and finishing point. In general it is used to describe overall performance, ie. how effectively the business unit has met its primary aim. Another important figure used in business is, *Return On Investment* (ROI), which is simply an indication of the benefit (eg. profit) that has been earned from a certain level of capital; in effect how efficiently the investment has been used.

It is necessary for this chapter to clarify the term *Budget* which has a slightly different meaning in education and business. Within education a budget generally refers to an amount of money (or perhaps hours) which has been allocated to a particular task or function. In business the term is used more widely and encompasses much of what is considered in education as targets as well as monitoring and control system[92]. In both education and business the measures used may be KPIs, as well as other high stakes proxy indicators.

Besides these measures which most businesses use[93], other business or industry specific figures may be reported, to show relative performance, and this may be in the form of league tables. For example, financial services companies will often show league tables of their investment performance relative to others (curiously they all appear to be the best – a lesson for schools!), car

---

[90] Some of their facilities management contracts include a bonus element based on exam passes.
[91] Such as measuring the difference between KS2 and KS3.
[92] Horngren and Foster (1991) define a budget as a: "...quantitative expression of a plan of action and an aid to coordination and implementation." (p. 5).
[93] Reporting profit is a legal requirement

manufactures may produce tables showing how good (for example speed or miles per gallon) their cars are compared to others; although probably the most objective are the football 'companies' whose results appear in league tables, and which in reality are probably very good indicators of relative performance.

Whilst much of the high stakes educational performance measurement systems (ie. KPIs) is a relatively recent phenomena[94], high stakes business accounting measures have been used for hundreds or thousands of years[95]. Furthermore, as Johnson and Kaplan (1991) state the main components of the modern accounting systems have also been in place for many years:

> By 1925 virtually all management accounting practices used today had been developed....These practices had evolved to serve the informational and control needs of the managers of increasingly complex and diverse organizations.
>
> (p. 12)

Although, many of the techniques had been developed by the early 20th century, there have been many significant developments in the use and application of accounts. This was in part due to the advent of the very large multi-national corporations, as well as the 'new' organisational forms (see chapter 4).

This chapter begins by looking at the work of three key theorists (Argyris, Hopwood and Otley) who in particular were concerned with the behavioural and potentially dysfunctional effects of budgeting (or target setting / reporting) systems. Their findings were not in complete accordance, therefore to help bring the various strands together, the work Birnberg *et al* is briefly considered. They provide a useful summary of the potential distortions to information systems, which may have dysfunctional effects.

Businesses like schools are increasingly using league tables to display the performance of different parts or sub-units of the overall organisation. One such company is TNT, and research by Moon and Fitzgerald (1996) is summarised to form a small case study on the organisational effects of league tables. This is followed by details of relatively unique experimental research by Keasey *et al* (2000) who looked at the behavioural effects of league tables. Over the years a number of theorists have asked the rather fundamental question as to whether accounting systems are in fact measuring the right things; the final part of this chapter looks at the work of Johnson and Kaplan (1991), who put forward arguments that management accounting has lost much of its relevance for modern organisations.

---

[94] With perhaps the notable exception of the Revised Code – See Chap 6
[95] Ezzamel (1994) traces the system of 'charge-discharge' back to ancient Egypt

Three Key theorists; Argyris, Hopwood and Otley

Argyris

Prior to the 1950's Birnberg *et al* (1983) point out that accounting and budgeting was considered a relatively straightforward technical process, and one that did not have any inherent behavioural implications. However, in 1952 Chris Argyris, an industrial psychologist published his seminal work, *The Impact of Budgets on People* (1952). This looked at the supervision systems in four production companies, and highlighted concerns about the possible dysfunctional effects of budgets.

Argyris found that the budgeting process could cause; an increase in job related tensions, unreasonable pressure being placed on subordinates, a fear of failure and the development of a blame culture within organisations, and in general a feeling of resentment and mistrust between superiors and subordinates. Such a culture could in turn lead to exploitation and defensive behaviour such as avoiding constructive criticism. Furthermore, 'accountants' and associated staff could become marginalized, and the actual construction of budgets might become something of a 'ritualistic' exercise, with game playing such as creating slack in the targets. Indeed, Argyris found that just the existence of budgetary controls themselves could provoke dysfunctional behaviour.

The 'solution' or way forward for Argyris was participation by all the relevant parties in the budgeting process. However, he also warned of the dangers of 'pseudo-participation', where the need for participation was emphasised by the senior management of the organisation, but in reality the line management merely resorted to imposing 'their' budget.

Argyris did recognise the significance and effect of the economic and business climate within which organisations existed. Factors such as these would significantly affect the feelings and responses of the employees. For example, a business that was thriving with perhaps all of the employees and their managers enjoying bonuses and job security, would probably display far fewer signs of dysfunctional behaviour.

Argyris work is commonly acknowledged as the first detailed study of the potential problems associated with budget systems. As Briers and Hirst (1990) point out his study was to provide an impetus for many other studies, as well as the increasingly widespread awareness and general acceptance (by many) of the potential effect of budget systems on behaviour;

> The Argyris (1952) study provided a milestone in behavioural accounting research. It went beyond simply examining alternative uses of accounting information in performance evaluation. Importantly, it was a comprehensive study, rich in theoretical, as well as empirical, content. Argyris was not only

> interested in the antecedents and outcomes of supervisory style, but also suggested budget participation as a possible moderator of supervisory style. Consequently, Argyris provided a broad foundation and, indeed, motivation for future studies.
>
> (p. 376)

Hopwood

Among the many studies that followed Argyris, quantitative research by Anthony Hopwood (1972) has proved particularly significant. This looked at a very large US steel company and although Hopwood looked more specifically at supervisory style, rather than budget pressure, much of his work built on and further developed Argyris's. In broad terms it did support Argyris's main findings (Briers and Hirst 1990).

Of particular interest to this thesis, Hopwood identified three supervisory styles; *Budget Constrained, Profit Conscious and Non-Accounting*. In the *budget constrained* model managers and business units are rigidly evaluated in terms of their detailed budgets. The equivalent in education would be, assessing teachers in terms of meeting their individual targets and their contribution to the KPIs. The *profit conscious* model was used to describe situations were managers were aware of, or / and worked towards an overall indication of performance which was related to their primary aims, in effect for a 'profit organisation' the overall profit or for a Strategic Business Unit (SBU) their contribution to this figure. Within education this equates to broader measures such as the overall 'value added' by the school. *Non-Accounting* measures include non-financial measures, for example new product development times or customer satisfaction ratings. Within schools attitudinal indicators, such as those within the YELLIS system would be similar.

Table 6 from Macintosh (1985) provides a summary of some of the key findings from Hopwood's (1972) research. The budget constrained model is closely related to, and to a large extent based on, the issues considered by Argyris's research in 1952. The results are reported here to aid and illustrate the discussion about the general principles, rather than trying to make detailed comparisons with the educational KPIs. That said, as will be seen later in this thesis, a number of these questions have been re-formulated and applied to the detailed research questions.

Table 6: Hopwood's style of evaluation

	STYLE OF EVALUATION		
	Budget Constrained	Profit Conscious	Non Accounting
(1) Relations with supervisor			
(a) Trust	3.3	4.2	4.0
(b) Respect	3.7	4.6	4.3
(c) Reasonableness of expectations	3.3	4.1	3.9
(d) Satisfaction	3.3	4.6	4.4

(2) Relations with peers			
(a) Supportiveness	3.6	3.8	3.9
(b) Agreement	3.4	3.7	3.8
(c) Helpfulness	3.8	4.0	4.2
(d) Friendship	3.3	3.7	3.7
(3) Rating of evaluative criteria			
(a) Cost concern	4.3	4.8	4.1
(b) Effort	4.1	4.7	4.5
(c) Quality concern	4.5	4.7	4.5
(d) Meeting budget	4.4	4.5	3.9
(e) Attitude to work and firm	4.1	4.8	4.6
(f) Co-operation with colleagues	3.7	4.3	4.4
	(N = 33)	(N = 43)	(N = 73)

Indices reported as means of responses on 1-5 Likert scales
From Macintosh (1985: 18) based on Hopwood (1973)

In terms of relationships between employees and supervisors the results suggest that the budget constrained model is more dysfunctional than the profit or non accounting systems, and this does of course very much support Argyris's findings. Interestingly there appears to be little difference in terms of peer relationships between the different systems. However, both Hopwood and Argyris suggested that an emphasis on budget based measures could encourage subordinates to 'gang together' which in its self may well be dysfunctional; there is a fine line between mutual co-operation and mutual resistance. Both Argyris and Hopwood gave possible examples such as: resisting the management's push for higher performance, manipulating results and even adopting 'wrecking' behaviour. The final group of questions suggested that subordinates like working towards broad quantitative targets rather than detailed budgets, or perhaps surprisingly non – accounting measures.

Besides looking at supervisory style Hopwood (1972) also argued that accounting performance measures were by definition only partial representations of actual or overall performance; ie. proxy measures. In addition, they may represent performance that is not necessarily controllable or even attributable to those being judged, and have a tendency to emphasise the short term (ie. what ever the accounting period under consideration – typically a year). As pointed out before, educational KPIs are by definition proxy indicators, they may be used to judge people who have no direct influence on the particular aspect of performance, and they may too encourage short-termism.

Otley

Subsequent research by David Otley (1978) which looked at individual operating units in the National Coal Board (NCB) aimed to replicate and extend Hopwood's work; however, Otley did

not reach the same conclusions. Although, he found a link between supervisory style and the nature of the relationships in the organisation, he did not find significant evidence to support the notion that a budget constrained[96] style of management necessarily had dysfunctional effects. Otley argued that such behaviour was primarily the result of how a manager felt about his (her) targets, rather than the actual supervisory style. He did nevertheless find that a budget constrained style encouraged practices such as 'income-smoothing'[97], bargaining for low or easily attainable targets, and other 'creative' accounting practices; which depending on how they are used my be dysfunctional.

In essence it was Otley's view that whilst budget systems did have an effect on behaviour (indeed this was one of the main reasons for using them) it did not follow that they would necessarily cause dysfunctional behaviour; much depended on the way they were used and the broader context within which they existed, or put another way their effect was contingent upon other situational factors. Macintosh (1985) provides a useful summary of Otley's views;

> Otley speculated that senior managers adopted an evaluation style which suited the prevailing circumstances and was, therefore, appropriate for each independent operating unit. These circumstances are dependent upon factors such as the toughness of the competitive environment, the general economic conditions, the size of the operating unit (and thus the relative magnitude of the investment in the unit), the relative experience of its manager, and the degree of profitability of the unit. Senior managers seemed able to match their evaluation style to each set of circumstances; for example, they exercised closer control through a budget-constrained style, on relatively inexperienced unit managers ....The managers generally felt that a certain degree of emphasis on meeting budget levels is helpful for technical efficiency and interpersonal relations....Prevailing circumstances may have more influence on the use of accounting information than does the individual's own personality and philosophy of management.
>
> These results point toward the need to develop a more contingent theory of budgeting control based on differences in organizational types, the environmental circumstances in which they operate, and the norms and values current both within the organization itself and within the society in which it is set.
> (p. 23)

Besides these three key theorists, a great many other research projects have not been able to unambiguously answer the crucial question as to the degree to which accounting measures may have dysfunctional effects. Indeed, a meta analysis by Shields and Shields (1998) of 47 studies provides little in the way of clear-cut or straightforward answers. Therefore many theorists have used a 'contingency[98] theory' approach. However, as Otley (1980) points out; "Contingency theories of management accounting have become a current vogue but have produced few significant results". (p. 413)[99]. The key limitation of contingency approaches is in not being able to

---

[96] Otley rejected the discrete classification of management styles, and replaced it with a continuum
[97] Eg. Carrying over income from a good month to a poor month
[98] Contingency theory simply means basing the theory on the situation; it can also be called situational theory
[99] It should be noted that this statement was with made reference to a number of arguments as to how the contingency approach to accounting could be improved.

necessarily draw valid inferences or make generalisations across different sectors and organisations.

From the perspective of organisational behaviour, Birnberg *et al* (1983) point out the majority of the studies focused on individuals rather than groups;

> The research basically has taken a micro orientation dealing with the motivational and psychological aspects of budgets focusing on individuals rather than on groups and on personality rather than on process. Its major contributions have often been diluted by conflicting results.
>
> (p. 121)

Furthermore, as many authors (eg. Van der Stede 2000, Otley and Fakiolas 2000) have pointed out there have over the years been many difference in the classification and terminology between studies. For example, Shield and Shield (1998);

> Future research would be more valuable if it were to provide explicit links between a study's nomological network and those of other studies to facilitate the development of general theories ...[of] management accounting variables.
>
> (p. 67)

The concept of making like with like comparisons has bedevilled much of the research. For example, Hopwood (1972) studied a large US steel corporation, whereas Otley (1978) studied the British National Coal Board. There would have been many significant differences in these two organisations which would have been difficult to control for. Indeed the time difference, with the 'oil crisis' occurring between the two may have had an effect, as might the prospect of de-nationalisation for the NCB.

Potential distortions to information systems

In spite of the difficulties of making firm predictions about the effects of budget and accounting systems, over the years various potential distortions to information systems, which may cause dysfunctional behaviour, have been identified and acknowledged. Birnberg *et al* (1983)[100] provide a useful summary based on six broad categories; *Smoothing, Biasing, Focusing, Gaming, Filtering, 'Illegal Acts'*.

Smoothing – *occurs when the manager is able to affect the natural or pre-planned flow of data without altering the actual activities of the organization.*

---

[100] Statements in italics taken from Birnberg *et al* (1983: 119-124)

Typically businesses may do this to bring forward a sales invoice to help meet that month's budget. Such behaviour may also occur in other parts of the public sector, for example to help a surgeon meet 'his' target for the month. The opportunity for such behaviour in education is less obvious, in terms of KPIs exam passes are only counted at the 'correct' age for the child. The official presentation of data is in a set format, although the media and schools themselves may be able to smooth the data to give a more favourable impression in their own publications[101].

Biasing – *those situations where the manager selects from a set of possible messages a signal that is likely to be accepted and is most favourable to him.*

Such situations usually exist when managers are able to influence their own targets, and build in slack. This will generally occur in situations of information asymmetry; for example a sales manager may know of a forthcoming contract and be able to use that as 'insurance' on meeting his target. In a similar way a school with detailed knowledge of the potential of its pupils may be able to agree an 'easy' target with the LEA.

Focusing – *occurs when certain aspects of the information set are either enhanced (highlighted) or degraded (hidden).*

If the opportunity exists most people or organisations will try to accentuate the best aspects of their performance, and keep quiet about weaker ones. A company's annual report contains both legally required information, and other more subjective information; and a manager, just as with his counterpart in a school, will tend to accentuate the best aspects. In the case of KPIs a school may emphasise its total point score, as opposed to the percentage of 5 A-Cs, or vice versa.

Gaming - *those behaviours where the sender through his job related acts causes the desirable message to be sent.*

Birnberg *et al* link this to the principle of using proxy indicators, where a manager may do a particular action which enhances what is measured. For example, delaying investment in new machinery, which will reduce that year's expenditure, and therefore increase the reported profit. Within education a similar result may be achieved by reducing the amount of teaching for non KPI subjects (eg. music and art), or avoiding 'harder' subjects.

---

[101] This may of course be a better way of presenting such data

Filtering *- the more desirable elements of the [data] are communicated and the less desirable ones are not (they also included delaying information in this section).*

A manager may be able to effectively filter performance information by a number of strategies such as, delaying, overloading and over aggregation (in effect loosing the important detail). Within education similar tactics can be used, in particular over aggregation. The overall percentage of 5 A-Cs will have the effect of filtering out the important components, such as the performance of individual departments and teachers.

'Illegal' acts *- are ones that violate private law (i.e. organizational rules) or a public law.*

Within a business this may include 'deliberately' mis-applying statutory accounting rules. For example, overstating the value of work in progress will in turn increase the reported profit. In schools recent examples include cheating to help children (schools?) perform better in the KS2 tests.

Birnberg *et al* point that there can be a significant degree of overlap between the groups, for example smoothing can be used as a form of filtering. They also acknowledge that the individual behaviours may not necessarily be dysfunctional;

> In response to Professor Buckley's comments we note that these behaviours may not always be dysfunctional. However, in our opinion the potential for dysfunctional effects to the firm is greater than the potential for beneficial effects.
>
> (p. 119)

Much depends on the particular circumstances. For example, within education it can be argued that concentrating on literacy to the partial exclusion of other subjects, may not be dysfunctional, in that it may allow those students to access other parts of the curriculum. However, if this were done primarily as a means of achieving KPI targets, that would seem dysfunctional: much depends on the underlying motives.

League tables in business – A case study

This section looks more specifically at the use of rankings or league tables in business. Many organisations use some form of rankings; for example, as part of employee appraisals or to display the performance of individual operating units. In terms of employee appraisals, Enron the discredited energy company, used the infamous 'Rank or Yank' to evaluate individual performance. Those in the top 5% received substantial bonuses, and those in the bottom 5% were sacked. Not surprisingly this encouraged individuals to engage in dysfunctional behaviour, which

ultimately contributed to the company's downfall[102] (Shaw 2002). More relevant however to this thesis is the use of league tables to report the performance of individual parts or profit centres of large organisations. TNT the delivery company provides a good example of such an organisation, and the discussion below is based on Moon and Fitzgerald's (1996) paper. *Delivering the goods at TNT: the role of the performance measurement system.*

The Indicator / Reporting System

TNT worldwide is a large (100,000 employees) and relatively diverse organisation, and a market leader in what is a competitive and expanding industry. However, this study is concerned with the organisation and control of the UK part of the company, *TNT Express (UK) Ltd.* The company is organised around a hub structure with a central headquarters and 28 regional depots. The depots, headed by a general manager, are profit centres, in that they are responsible for generating income. Nevertheless, there is a high degree of central control, and structurally the system has a number of similarities with the education system.

In terms of performance information, raw data is fed up to head office and then various reports are distributed back to the depots, although some performance indicators are monitored on a continuous or real time basis; for example, meeting delivery times. The use of league tables to 'drive up' performance is central to the philosophy of the company:

> A central feature of the performance measurement system at *TNT* is the widespread use of league tables that display each depot's performance relative to one another. These emphasize the company's critical success factors of profitability and quality of service by reporting results on these dimensions at the depot level weekly. Implicitly, competition, in terms of performance, is actively encouraged between the depots. An individual's [depot] position in the league table is keenly observed both by that individual and his (her) peers.
>
> (Moon and Fitzgerald 1996: 454)

A total of four reports are produced and distributed on a weekly basis. The main report shows the profit (absolute.and as a % of revenue) of each depot relative to other depots. Cumulative figures are also given to help smooth the data. Reports are also produced for sales and customers care, deliveries, and finance and administration. The sales and customer care report looks both at sales targets and customer satisfaction. The deliveries report ranks depots in terms of factors such as punctuality and mis-delivery, and the finance and administration report looks at how efficiently the business is run; for example, the length of debtor period. All of these reports and their associated tables are high stakes, in that substantial bonuses and potentially disciplinary action may follow. Furthermore, these reports and tables are, to a degree, used as proxy indicators by the head office.

---

[102] NB: The key factors were the accounting methods and failure and / or complicity of the auditors

The financial report (net profit) is the key indicator for depot performance, with the other reports aiming to give a broader picture in terms of current and potential future performance. In addition, it is claimed that measuring these other aspects of the business will discourage dysfunctional behaviour such as maximising profit at the expense of other important objectives and functions.

Some of the relevant surrounding issues from Moon and Fitzgerald are discussed below, and where relevant related to the education system:

Transmission of corporate values

The performance management system is explicitly used as a mechanism to transmit the corporate values throughout the company;

> [The performance measurement system has] enabled the company to drive down the corporate message that the business needs to be profitable, there is no market for excuses and, in an organization whose product is essentially the same throughout the country, what can be achieved in one region ought to be replicable elsewhere. Issues of controllability or equity are seen as irrelevancies in the context of this broader corporate drive.
>
> (ibid: 455)

There are echoes of a zero-tolerance philosophy that will be familiar to many in the public sector, as will the argument that if one school achieves good results in one area, then there is no reason why others shouldn't as well. The management philosophy at TNT is very much to 'drive' the company and its employees hard, in part because this is how the industry as a whole operates.

Pressure and stress

Linked to the above, one observed effect of the system has been to increase the pressure on individuals to work increasingly harder;

> .... most depot personnel interviewed were aware that their own workloads had increased substantially over the last few years. The volume of consignments collected and delivered had increased at a faster rate than staffing levels. Long hours seemed to be becoming the norm for many, and stress levels were higher. While the performance measurement system was not directly responsible for this, it was seen as implicated to some extent, as it provided the data for head office benchmarking.....
>
> (ibid: 450)

This may well be of benefit to customers with lower charges, however there of potentially negative consequences on the employees which may rebound on the company and society as a whole; for example, increased health and benefit costs, as well as more intangible social costs. To an extent this is indicative of the market that they are in, which is in part fuelled by the potential further privatisation of the postal service in the UK. Similar arguments about increased stress and workload because of league tables, have been made in many areas of the education system; for example, the now semi-deregulated further education sector, which is operating in an increasingly competitive market. This is however not to suggest that all stress is bad, it is commonly accepted

that some degree of stress can be beneficial in terms of performance, although the 'type' of stress (ie. how it is imposed and the degree of control the individual has) is also significant.

Control and responsibility

Although depots are individual profit centre's they are nevertheless affected by the actions and performance of other depots and parts of the company;

> ....depot managers and depot personnel are held responsible for areas over which they have no formal control. The network nature of the business implies that there is a high interdependence of depots: the collecting depot will not necessarily be the delivering depot. Business may be generated for which the collecting depot receives the revenue, which is in fact difficult to deliver, but the delivering depot bears the cost. This impacts on both depots' profit statements.
>
> (ibid: 454)

In education too it can be difficult to assign responsibility for the eventual output performance. For example, a year 11 geography teacher's apparent performance will be influenced by other teachers 'further down the line'; indeed not just those teaching geography, but those teaching cross functional skills, such as drawing graphs and comprehension.

Equity

A primary aim of any performance measurement system is to provide a fair and equitable indication of performance by measuring like with like. However, some depots have significant inbuilt advantages over others;

> In theory, performance is transparent. In practice, although each depot performs essentially the same function and is measured in the same way, their circumstances may be very different. Some may be near to the hub, some may be far away; some may be located in areas with high collections, some in areas with high deliveries, some in urban zones with well developed road networks, some in remote, rural places. Measuring performance via the league tables makes no attempt to allow for these relative differences. Inequity is built into the system.
>
> (ibid: 445)

Interestingly the paper gave the impression that this was accepted within the organisation and did not appear to be a particular bone of contention. A similar issue exists with school league tables, where the location (and therefore intake) will have a significant bearing on the relative performance of individual schools.

Informal networks

Although the organisation is tightly controlled by the head office at the centre, and competition is encouraged between depots, and therefore by definition one depots loss is another's gain, there was nevertheless evidence of inter-depot co-operation. For example, in terms of delivery problems:

> .... the delivering depot will discuss the problems informally with the collecting depot. These informal discussions are facilitated by the close communications between depots recognizing the interdependencies of the business. These findings are against the conventional wisdom of

> management accounting texts but in agreement with the field studies of Merchant (1987) and Otley (1990).
>
> (ibid: 454)

In a similar way in education one school's loss in the league tables is its neighbour's gain. And as with TNT informal networks, for example cluster meetings, may help ameliorate the potential dysfunctional effects.

Although parcel delivery and education are clearly very different businesses: TNT simply takes a package from A to B in a day or so, whereas the education system takes a child to adulthood over a period of many years. TNT's indicator system does nevertheless exemplify the difficulties and complexities of measuring performance, as well as the potential dysfunctional effects. They have attempted to offset these by using (to an extent) a balanced range of indicators; indeed, given the developments in education there may well be some useful lessons that can be learnt. Moon and Fitzgerald (1996) summarize the difficulties and challenges faced by TNT in the development of their indicator system:

> It is almost an axiom of management accounting thought that organizations need to formally measure their performance. At one level this is straightforward. 'What is the bottom line? . . . by how much have sales grown this year? . . . how much profit have we made?' However, with the emergence of the global organization, the rapidity of technological advancement, and the hostility of competition, simple financial yardsticks, such as return on investment or residual income, are no longer sufficient as organizational performance measures, if indeed they ever were. Instead it has become increasingly important for organizations to develop systems of performance measurement that reflect this growing complexity of the business environment, and that monitor the organization's own strategic response to this complexity.
>
> (p. 432)

Experimental research on the effects of league tables

At the individual level there has been a substantial amount of experimental research on the effect of ratings for performance appraisals (Harris *et al* 1995), there has however been very little experimental research at the organisational level. Research by Keasey *et al* is relatively unique in this area, and the details below are taken from their paper (2000); *Performance measurement and the use of league tables: some experimental evidence of dysfunctional consequences.*

Keasey *et al* point to the increasing use of league tables both in the public and private sector, and much of their interest in the behavioural effects stems from the Financial Services consultation paper 28 (see Jones and Keasey 2000) on the likely introduction of *Financial Services League Tables*. A number of prominent scandals, such as 'pension mis-selling' have prompted concerns about the potential dysfunctional effects of high stakes proxy indicators, particularly when

presented in a league table format. Their research was largely concerned with, *measure fixation*[103]. (This particular issue was identified by Peter Smith (1995a) whose work is considered in the next chapter).

The experiment consisted of a number of participants[104] taking the role of branch managers of stores in a large national supermarket chain. Performance was assessed by the store's ranking in a league table, which was made up from a number of weighted components, including profit and other qualitative indicators such as absenteeism and customer satisfaction. As the experiment ran each 'manager' had to assess and consider different projects and proposals, which would either increase 'profit' and raise their position in the table, or the opposite.

From the point of view of this thesis there were two particularly significant findings. Firstly, it appeared that organisations which performed well; ie. had a high league table position were more likely to be innovative;

> ...league table position seems to have a bearing on the willingness of subjects to take on a given amount of risk but only where the initial position is high.
>
> (Keasey *et al* 2000: 281)

This would seem very relevant to education, suggesting that schools which are high in the tables are more likely to try new innovative approaches, and schools at the bottom may feel that they have to 'play safe'. However, it may be just those schools (with perhaps more 'difficult' children) that should be trying more innovative strategies. The message from government tends to be rather mixed, with on the one hand some schools being congratulated for finding new innovative approaches, and on the other, factors such as the NLNS and Ofsted inspection framework discouraging innovation.

Secondly, their research suggested that actions which would raise the organisation's position in the league took precedence over those which would simply increase profit[105];

> ...it seems, some subjects *simply* ignore the financial expectations (or at least decide the *downside* risk is sufficiently bounded) and choose on the basis of the given probabilities, the possibility of going top once more acting as a strong influencing factor. This *is interpreted* as clear evidence of dysfunctional behaviour, resulting from an over-emphasis of the impact on league table positions (measure fixation).
>
> (ibid: 282)

---

[103] Concentrating on measures rather than the underlying objectives.
[104] Students at the University of Leeds doing business studies courses. Keasey *et al* discuss at length the validity of using surrogates for such experiments.
[105] The 'managers' did not receive profit related pay.

This is particularly relevant to the possible developments in KPI systems. for example. many arguments have been made in favour of reporting value added. If these were reported in league tables, this may reduce the dysfunctional effects of concentrating on 5 A-C's or level 4s. however it may lead to other dysfunctional behaviour: (ie. those aimed at increase VA scores)[106].

In summary, Keasey *et al* emphasise the need for care not only in the use of league tables but also in the overall design of the performance measurement systems, including the actual measures themselves:

> In a business context it is no longer obvious that the use of league tables as a means of performance assessment and as a motivational device will guarantee meeting the entity's (private or public organisation) objectives. The implications of using league tables to assess and motivate sub-units of an entity are not well understood.....The results of this study clearly indicate that organisations must be cautious in the construction and use of such league tables in sub-unit performance assessment and motivation.
>
> (ibid: 285)

The Wrong Measures?

Alongside the research which considered the effects of performance indicators *per se* a number of theorists have asked what is perhaps the more fundamental question; 'are we measuring the right things', or put another way 'are the measures relevant to the business or organisation?' On the face of it, this question might appear naïve for a business; clearly 'for profit' organisations need to measure *profit*, and how much profit they make will be an indication of how they have performed. In a similar way, the return (ie. profit) on the investment (*ROI*) is likewise on the face of it a clear and relatively unambiguous indication of how efficiently the money has been used to achieve a certain level of performance. As previously mentioned these two indicators have very much come to dominate business performance reporting, although as Johnson and Kaplan (1991) argue, they are not necessarily relevant to the actual management:

> Today's management accounting information, driven by procedures and cycle of the organization's financial reporting system, is too late, too aggregated, and too distorted to be relevant for managers' planning and control decisions...And despite the considerable resources..... the figures do not measure the actual increase or decrease in value that has occurred during the period.
>
> (p. 1)

Johnson and Kaplan's (1991) seminal text, *Relevance Lost: The Rise and Fall of Management Accounting*, brought together many of the questions regarding the usefulness and relevance of these key accounting indicators. This work in particular has caused a lot of practitioners and theorists to

---

[106] See Dorn 1998 for a discussion of this in the US

question many of the fundamental accounting assumptions (Roslender 1996, and Drury and Tayles 1997).

Although the ROI ratio does not as such exist in education, it is closely related to output measures (ie. profit) which as previously argued is similar, at least conceptually, to the existing KPIs. Furthermore, there are signs that a ROI type of indicator will in the future become more important, and used as a high stakes indicator. In essence ROI aims to indicate how effectively inputs or resources are used, by showing the level of return (profit) for a given input. Within schools a similar figure can be calculated by finding the ratio between inputs (eg. ARPU[107]) and outputs (eg. % GCSEs). This already occurs in the independent sector, with for example the Sunday Times league table of best value schools (2001). Indeed the trend is increasingly for such indicators to be used, for example proposals from the government's (Treasury) Public Services Productivity Panel advocate; "A measure of relative value for money – that is to say, a measure of what is being achieved per pound spent per pupil in each school" (Mayo 2000: 10).

Relevance lost?

Profit and ROI have come to be used extensively throughout business as Johnson and Kaplan (1991:3) point out; "Financial measures such as ROI have become for many organisation the only measure of success". This is even though they have been widely criticised by many authors and theorists (Dearden (1987). Indeed, Swieringa and Weick (1987) make the point that ROI is criticised by most academics, to the extent where it is difficult to find a positive discussion of its use. Nevertheless, as Drury and Tayles (1997) point out, it is still very widely used (internally as well as externally), and on a practical level is valued by many managers.

Johnson and Kaplan identified three main areas of criticism of business accounting based on these (profit and ROI) measures; firstly, that management accounting had become subservient to financial accounting; secondly, that it failed to keep pace with the technological changes and the demands of industry; and finally, that the academic accounting world took too simplistic a view of organisations and developed overly crude theoretical models which did not reflect reality.

The first point, the relationship between management accounting and financial accounting is central to much of Johnson's and Kaplan's thinking. Management accounting is concerned with internal management information, and typically is used for planning and decision making. Financial accounting provides external performance information (primarily financial although other aspects may have to be accounted for), with much of the format being set out in legislation and

---

[107] Age Related Pupil Units – ie. the amount of money the school receives

'accounting standards'[108]. Broadly speaking educational 'accounting' systems are quite similar. Individual schools will have their own performance or management information systems, for example to help with planning and target setting, as well as the KPI system for external reporting.

Their other two points would also seem relevant to education. Many may well feel that the current indicator systems have not kept pace with technological changes: for example reporting the performance of new outputs (eg. NVQs and Modern Apprenticeships or assessing the performance of special needs teaching). And although much academic work has been carried out in terms of improving performance measurement techniques (eg. Value Added, reporting statistical variation and using a range of softer indicators[109]), this has to date, had a limited impact on the National KPI systems. Furthermore, the issue of relevance or 'suitability for purpose' of KPIs tends not to be a major source for debate in the field of educational management. A search of the NCSL website (including the 'knowledge base' produced no reference or acknowledgement of such questions). An approach to raise this issue on 'their' internet based network for school heads (Talking Heads) was rejected, in spite of initial enthusiasm. Therefore, the questions raised by Johnson and Kaplan may be pertinent to education;

> ....the academic literature concentrated on increasingly elegant and sophisticated approaches to analyzing costs for single-product, single-process firms while actual organizations attempted to manage with antiquated systems in settings that had little relationship to the simplified model researchers assumed for analytic and teaching convenience.
>
> (p. 14)

As a consequence of the claimed failure of management accounting, Johnson and Kaplan identified three main problems; firstly, the system fails to produce accurate product costs (in other words reliable and valid performance measures); secondly, the systems do not help improve performance (ie. reduce costs and improve productivity); and finally, and most importantly from the perspective of this thesis, they encourage dysfunctional behaviour.

Why Dysfunctional

Behind the apparent simplicity and transparency of *Profit* and *ROI* calculations, lies a potential 'can of worms' for many companies. Because of the way they are calculated, a company's performance can be made to appear to improve in both positive and negative ways. For example, it may sell more products or reduce wastage, which is obviously good; however, the same level of 'improvement' can be shown by reducing investment in areas such as research and development and training, and this kind of action may well have a negative long term effect. This point is further illustrated by Johnson and Kaplan (1991);

---

[108] These are set by the professional accounting bodies
[109] See Fitz-Gibbon (1996), Goldstein and Spiegelhalter (1999), and Audit Commission 2001

> ...perhaps the most damaging dysfunctional behavior induced by a preoccupation with short-term profit center performance is the incentive for senior managers to reduce expenditures on discretionary and intangible investments. When sluggish sales or escalating costs make near-term profit targets hard to achieve, managers often try to prop up short-term earnings by cutting expenditures on R&D, promotion, distribution, quality improvement, applications engineering, human resources, and customer relations-all of which, of course, are vital to a company's long-term performance. The immediate effect of such reductions is to boost reported profitability, but at the expense of sacrificing the company's long-term competitive position.
>
> (p. 201)

As well as potentially failing to properly indicate true costs or values, Johnson and Kaplan (1991) suggest that the accounting systems may allow or encourage 'creative accounting':

> Some of these [accounting] activities may create value to shareholders.... Still, it is hard to believe that a focus on creating wealth by clever financing and rearrangement of ownership claims will help companies survive in the global competition of the 1980s and beyond. Ultimately, wealth must be created by imaginative and intelligent investment in assets and the proper management of them, not by devising novel financing and ownership arrangements for assets.
>
> (p. 201)

There are many examples in the commercial world of such behaviour which have led to losses both for the company and its many stakeholders, and even broader society; Enron and Marconi[110] would be two recent examples. In the short term, such action may be of benefit to *some* shareholders, (ie. 'those in the know'), as well as employees whose remuneration is determined by the dependant profit related measures. Furthermore, there are incentives for other bodies, such as auditors, investment advisors and even the media to collude with the illusion.

Within education there may be opportunities for schools to reduce their investment in long term performance to help meet short term targets, for example, booster classes; although there is far less opportunity for 'creative accounting' in terms of the actual teaching processes. However, the increasing convergence of performance and accounting systems, due in part to the growing involvement of 'private sector' companies, and the consequential 'need' to incentivize individuals may make this a more relevant issue in the future.

Increasingly, for many organisations, including education, there is a trend towards leaner more decentralised organisational structures, which in effect allow the centre to manage the branches at a distance[111]. Johnson and Kaplan, and indeed many others, warn of the dangers of using simplistic high stakes proxy indicators (in particular profit and ROI) for this purpose. They warn of the dangers of decisions being made on numbers being fed up the system:

---

[110] See Shaw (2002) and Randall (2001) for a discussions on these two companies
[111] See *principle agency* (PA) theory chapter 2

> Financial managers relying exclusively on periodic financial statements for their view of the firm, become isolated from the real value creating operations of the organization.
>
> (ibid. 31)

In a similar vein, Munro (1995) highlights the problems of de-contextualizing performance information;

> Managing by distance takes the artefact of accounting numbers to its logical conclusion... a dissemination of accountability away from the line, through 'output' measures which individuate and intensify responsibilities, facilitates a switch to an *insistence* on a number being *met*, at all costs to the individuals concerned. In this way, a propensity to give 'accounts' which excuse, legitimate or justify their failures is silenced.
>
> (p. 146)

There is evidence of this approach being increasingly adopted in education. The trend is towards KPI targets being set centrally (Hacket 2000), which combined with recent comments by the last education secretary, suggests that ever greater importance will be placed on the numbers (KPIs) fed up the system;

> I know at the touch of a button the performance of every single one of my 24,000 schools. I know that it is my responsibility to spot failure before it happens,... I do not think it could have happened four years ago because the accountability mechanism were not in place.
>
> (Morris 2001b: Q2).

And as well as being used for 'management' purposes, for example, macro resource allocation, the results are being used increasingly to determine both rewards (eg. school achievement awards) and punishment (a visit from Ofsted).

In essence Johnson and Kaplan (1991) argued that the main problems with the existing key indicators (profit / ROI) is that they do not necessarily measure the aspects of the organisation which actually contribute to its long term performance;

> The opportunity for companies, or their profit centers, to increase reported incomes by sacrificing long-term economic health illustrates a fundamental flaw in the financial accounting model. The flaw compromises the role of short-term profits as a valid and reliable indicator of a company's economic health.
>
> (p. 200)

Furthermore, they argue that simply combining indicators or producing ratios does not necessarily reflect the true value of the organisation, indeed any more than a simple summation and ranking of GCSE performance for a school;

> A company's economic value is not merely the sum of the values of its tangible assets, whether measured at historic cost, replacement cost, or current market prices. It also includes the value of intangible assets: the stock of innovative products, the knowledge of flexible and high-quality production processes, employee talent and morale, customer loyalty and product awareness, reliable suppliers, efficient distribution network, and the like.

(ibid: 200)

The Ascendance of Profit and ROI

Although it is relatively easy (certainly at the theoretical level) to criticise the indicators, it is useful to be able to understand how they have evolved and developed. As previously discussed in chapter 4, measuring performance, and by definition the need for accountability, was central to the organisational systems that evolved at the turn of the century, such as scientific management.

Hoskin (1996) however, adds an interesting perspective by arguing that the 'need' for accountability in businesses largely originates from the education system, in particular the formal examination systems that developed throughout Europe in the second half of the eighteenth century. He cites, for example, the *Tripos, Extempore and Seminar*;

> This, I suggest, is the context in which the 'awful idea of accountability' could suddenly emerge and become so irresistible. This transformation, whereby the pedagogic arena suddenly became flooded by numbers combining 'is' and 'ought', and wherein we became selves whose rational identity was constructed through a mix of external and internal examination, produced a context where 'accountability', to self and to others, could become a self-evident rational good.
>
> (p. 269)

Although the principle and concept of profit itself goes back thousands of years, ROI was developed in the 1920's, (as with many of the management systems in that era, by an engineer). Its primary use was for the management and control of the new very large multi-divisional firms, such as Du Pont and later General Motors. Significantly, ROI reports were intended only to be used by the top management;

> It would be a mistake, however, to conclude from their extensive use of formulae and charts that the founders of the Powder Company "ran the company by the numbers." The return on investment reports went only to top management who used the information for planning and company-wide control. Subordinate managers were not compelled to achieve return on investment targets; rather, they strove to achieve economies and efficiencies within their respective specialized activities. Top management alone assumed responsibility for the investment and allocation decisions that ultimately determined how effectively the company's integrated activities used capital.
>
> (Johnson and Kaplan 1991: 84)

A far cry from today's high pressure accountability with 'league tables' and 'Name and Shame' holding individuals at all levels responsible for the organisations KPIs: whether this be in terms of profit or exam passes.

Johnson and Kaplan (1991) point out that much of the driving force behind these indicators came from the finance and investment community. This includes investment analysts, and auditors whose professional duties required them to make safe, objective and verifiable statements. Such people

have to make judgements on many different organisations, and as human beings there are practical limitations on the number of different indicators they can comprehend: for many just one overall 'magic' figure is ideal. In this context, Ray Stata (CEO of Analog Devices), discusses the problem of dealing with large amounts of conflicting information, and argues that in the final analysis, the messages from the indicators which contribute to the 'bottom line' (ie. profit) will prevail; "When conflict arise, financial considerations win out" Stata (1989:63). Furthermore, Wilson and Chua (1993) make a number of pertinent observations on these conflicting pressures;

> Whatever index of performance is chosen must encompass a variety of factors, but it is extremely difficult to conceive of a single measure that embraces all the key factors. . . An index based on cost is likely to be too narrow since it excludes revenue issues. On the other hand, a profit indicator accommodates both revenue and cost but excludes such important issues as customer loyalty; (ii) establishing good employee relations; (iii) building a skilled team.
>
> (p. 364)

A further reason for the ascendance of profit and ROI is the key part they have come to play in controlling organisations;

> Until the 1920s, managers invariably relied on information about the underlying processes, transactions, and events that produce financial numbers. By the 1960's and 1970s, however, managers commonly relied on the financial numbers alone.
>
> (Johnson and Kaplan 1991:125)

Related to this is the change in the type and backgrounds of those managing the organisation;

> Early twentieth-century organizations such as Du Pont, General Motors, and General Electric had been created by owners who understood the technology of their products and processes. In succeeding decades, however, chief executives were selected whose entire careers had been spent in staff functions such as accounting, finance, and legal. Lacking knowledge in their organizations' underlying technology, executives increasingly made decisions based upon their projected impact on short-term financial measures, especially earnings per share and return on investment.
>
> (p. 14)

This would seem to go to the heart of much of the current debate of who runs the education system; the teaching profession or the government (See Baker 1994). And further into the area of public accountability and trust (see O' Neill 2002).

Criticism of Johnson and Kaplan

Whilst Johnson and Kaplan's work has had a significant impact on the thinking of both academics and practitioners, a number of their arguments have not been universally accepted; indeed such questioning may well be indicative of the healthy debates that occur in this field. As with much of 'Guru Speak' whilst it is intuitively very appealing, and undoubtedly very sensible, it is far more difficult to 'prove' the points and quantify the effects; an issue considered by Drury and Tayles (1997);

> Johnson and Kaplan's claim [that management accounting is becoming subservient to the demands of financial accounting] is based mainly on anecdotal evidence and little research has been conducted on whether management accounting information differs from that which is necessary for meeting external financial reporting requirements.
>
> (p.263)

Drury and Tayles themselves carried out research on over 300 companies. The results, however, were relatively inconclusive;

> The survey findings indicate that many companies use the same information for both internal and external reporting, even when conventional wisdom suggests otherwise. We conclude from the findings that Johnson and Kaplan's claims that financial accounting dominates management accounting cannot be rejected and there is a need for further research.... The fact that information that is produced for management accounting is not inherently different from that which is required for external reporting is not sufficient evidence by itself to indicate that management accounting is subservient to financial accounting.
>
> (ibid: 272)

Furthermore, it is worth noting that they found a significant number of practising managers supported ROI, having found it to be a useful incentive mechanism. In a similar way before this, Burns *et al* (1994) undertook a survey of the opinions of UK management accountants and concluded that there was little evidence that external reporting dominates internal reporting. However, they do suggest that the responses may reflect the degree to which external reporting has become integrated with internal reporting systems. In other words the measures that the managers had 'grown up' with, and knew best. It may also be the case that they are most familiar with the limitations of these measures. Furthermore, Hopper *et al* 1992) argued that the financial accounting measures have become so ingrained in the business organisation culture, to the degree, that the deficiencies may be hidden.

Like Burns *et al* it should be pointed out that Drury and Tayles' research was aimed at professional accountants who had been qualified for more than five years, and were employed in medium to large (turnover exceeding £10 million) organisations. It can be argued that these are the people most likely to support these indicators, other managers; for example, production, sales or human resources, might take a different view.

Ezzamel *et al* (1990) question both Johnson and Kaplan's account of the history of management accounting, and their remedies. They take a different view of the history and argue that using the 'numbers' to manage is an essential part of modern organisations.

Indeed, Johnson's (1994) own views have to an extent changed, and he would appear to be less salient about the technical changes to the indicators which he had previously advocated;

> Relevance was not lost by using improper accounting information to manage, it was lost by improperly using accounting information to manage.
>
> (p. 262)

As Roslender (1996) points out; "Johnson now sees management accounting as part of the problem...rather than providing a solution to it". (p. 545). In effect, the argument has moved on to question the use of much of the management accounting systems. Various improvements have been advocated, for example the use of a balanced range of indicators (ie. Balanced Score Cards – which is discussed in chapter 14). However, for an increasing number of companies, for example Asea Brown Bover (ABB) and Unilever, the solution is to move beyond the traditional budget system, and break through the barrier to 'the third wave'. In essence this requires a completely new decentralised network structure, with very different types of accountability, which maximise the intellectual assets of the organisations. In effect using the accounting systems to help build new structural forms that will allow and encourage the development of the learning organisation (See Hope and Fraser 1997 and www.beyondbudgeting.org).

Conclusion

This chapter has taken a broad view of the underlying theories of high stakes business performance indicator systems, and in particular the measures of profit and ROI. It has highlighted the difficulties faced by the theorists in this field of making valid like for like comparisons, due to the significant structural and contextual differences between organisations. It has also been argued that similar challenges and problems face those using and researching the effects of educational KPIs. Nevertheless, although important differences remain it is posited that the increasing convergence between the sectors should allow greater learning, *both ways*. And based on the theories and experiences discussed in this chapter, a number will be used in the detailed research questions later in this thesis.

A further important message from this chapter relates to the development and evolution, over many years, of accounting theory. This field critically and independently considers both the behavioural aspects of using accounting measures, as well as the appropriateness of the measures themselves; furthermore, it does tend to recognise the inherent complexities of organisational systems and the effect of contexts within which they exist. This is however not to suggest that the field is in agreement, far from it, indeed much of the learning comes from the many different views and interpretation. Birnberg *et al* (1983) provide a useful illustration:

> Accounting research in the area of information and control systems has gradually evolved from a focus on budgetary control to a broader organizational view. This evolution results primarily from a

> changing perspective of organizations, one that is now characterized by a richer, more detailed understanding of the underlying organizational processes.
>
> Within accounting research, this changing view has led to a questioning of prior approaches taken to the study of accounting as an information control oriented discipline. More specifically, it is now recognized that prior accounting research has been too narrowly defined and as a result has ignored many of the realities of organizational functioning.
>
> (p. 111)

This is largely in contrast to education, where KPI systems and the associated management thinking is still very much in its infancy. At present, KPIs appear to be viewed as unquestionable third party 'add-ons' to the educational management process. A recent question on a professional accountancy exam paper asked candidates to critically evaluate; *In recent times, the traditional annual budget has been accused of being incapable of meeting managerial demands in a competitive environment.* It may be a while before candidates sitting the NPQH are asked the educational equivalent.

Chapter 7

High stakes indicators and other public sector organisations

The performance of caring professions, such as teaching, can be inversely proportional to the quantifiable outputs (Gorz)

Introduction

This chapter looks at some of the evidence and theories relating to possible dysfunctional consequences of using high stakes performance indicators in the public sector. There has been much anecdotal evidence and practical examples of such behaviour, although little in terms of specific research (Jacobs and Manzi 2000). It is suggested that much can be learnt from the experiences of other public sector organisations, because like education, there tends to be complex and sometimes conflicting, as well as intangible outputs and outcomes, which PI systems are attempting to quantify. Other areas of the public sector have many contextual similarities to education; for example, government accountability and control structures such as Ofsted (and its equivalents) and PSA targets. Finally, as with schools, PIs are also being used in many public sector organisations to 'drive through' changes and the government's reform agenda.

Unlike the business accounting field there has not been a long history of critical research on the effects of PI systems, although increasingly theorists and practitioners are considering the behavioural implications. The much cited work of Peter Smith provides a good summary of some the ways in which PIs can have unintended consequences. This work is considered first, and is followed by brief details of three other areas of the public sector; *Health*, *Rail* and *Police*.

Smith – Unintended consequences of PI systems

Peter Smith's (1995a) paper: *On the unintended consequences of publishing performance data in the public sector*, identified eight / nine[112] potentially dysfunctional effects of publishing performance data in the public sector; *tunnel vision, sub-optimisation, measure fixation, myopia, complacency, misrepresentation, gaming, misinterpretation, ossification*. This was in a similar vein to the six factors identified by Birnberg *et al* (1983) which was considered in the last chapter. Fitz-Gibbon (1997) has shown this model to be very relevant to education. She used it to assess the possible consequences of publishing value added data, and found that with one exception, all of the points had been commented on by headteachers in response to open-ended questions. The model is discussed below with reference to Smith (1995a) and Goddard and Smith (2001).

Tunnel vision - *Concentration on areas included in the performance indicator scheme to the exclusion of other important unmeasured areas.*

---

[112] He originally identified eight factors, but this has increased to nine.

This is very much at the heart of the issues considered by this thesis. In effect the danger of schools concentrating on those items measured by KPIs: for example, English primary schools 'over' concentrating on the subjects tested by the KS2 tests, and excluding other such as art and sport.

To counter this effect Smith, like many others, advocated the use of a wide range of indicators, for example balanced score cards (see chapter 14).

Sub-optimisation - *The pursuit of narrow local objectives by staff at the expense of the objectives of the organisation as a whole.*

This can occur at various levels in the system. For example, at the National level lack of physical education may have a negative effect on health service resources. LEA wide inclusion targets, may be viewed as sub-optimisation by the school receiving the children. And at the school level a department or teacher getting a disproportionate amount of resources (or time) for 'their' subject.

Smith emphasizes the use of joint responsibility in planning and setting targets to help counter this effect. For example, in the case of children's health at the national level, this could include the sports council, the NHS as well as the DfES.

Measure fixation - *Pursuit of success as measured rather than as intended*

As discussed with reference to Keasey *et al* in the last chapter, an organisation may concentrate on its league table position, rather than its overall performance. For example, schools concentrating on the percentage of 5 A* to Cs, (on which league tables are based), rather than the overall quality of passes.

Smith advocates the greater use of 'front line' staff in the setting of targets. Within education such effects might be countered by setting individual targets for children, rather than overall school or LEA targets.

Myopia - *Concentration on short-term issues to the exclusion of long term considerations which may only show up in performance measures in many years time.*

Education is by definition a long term process, and many skills, for example languages are best learnt at a young age. However, pressure to meet KS2 targets may discourage primary schools from this, and the consequences may not be evident until many years later.

The use of process or intermediate targets may help ameliorate these influences. For example, in the case of language, schools having targets based on individual pupils throughout the education process, rather than at the output stage (eg. GCSE)[113].

Complacency *- Lack of ambition for improvement brought about by adequate comparative performance*

Although politically a rather emotive subject, given the many claims by the government that the public sector breeds complacency and therefore needs to be reformed (eg. Webster and Sherman 1999), Smith is concerned with how the measures, rather than the macro politics may contribute to this behaviour. He gives the example of individual managers preferring to stay out of the headlines and be perceived as OK, rather than risk being 'too obvious'. Other social and contextual factors, besides the indicator systems, may well contribute to such behaviour (see for example the Hawthorne Studies chapter 4).

Smith warns against measurement systems which have threshold performance levels, for example in the health sector the traffic lights system[114]. Targets which aim for continuous improvement may help discourage complacency.

Misrepresentation *- The deliberate manipulation of data by staff – ranging from 'creative' accounting to fraud - so that reported behaviour differs from actual behaviour*

At the extreme end of the scale this would include cheating to improve the figures. This might be in the overtly dishonest form of helping a student with a test, or perhaps less so by excluding weak (and badly behaved) students. Behaviour such as taking a creative view of attendance may also be a misrepresentation of the true figures.

Increased auditing and checking is suggested as a way to prevent this, although the underlying contextual motives which cause such behaviour may be more pertinent.

Gaming *- Altering behaviour in order to obtain strategic advantage*

Typically this involves setting easily achievable targets. Fitz-Gibbon (1996) gives the example of primary schools being as severe as possible in the base-line assessments, which will increase the

---

[113] Although for languages there are no National KPI targets

[114] The tables band trusts in terms of red, amber and green (see DOH 2002)

apparent gain at KS2. In addition, deliberately under performance one year, may allow easier targets to set the following year.

A possible solution is to use some form of benchmarking. Results from third party non-high stakes indicator systems could be used to help set targets; although again the underlying context is probably a significant factor.

Misinterpretation - *Incorrect inferences about performance brought about by the difficulty of accounting for the full range of potential influences on a performance measurement*

Smith gives the example of not properly allowing for case mix in interpreting results. Within schools different cohorts will perform at different levels, in particular if they are small. Therefore, results will not simply improve year on year, but will vary, and this variation may well in reality not be significant. Fitz-Gibbon points to the confusion surrounding 'correlation and causation', which may result in the wrong conclusions being drawn. For example, a school may adopt a new policy regarding truancy, and truancy may go down; however the two may not be linked. The reduction may be as the result of some other factors, and because schools are such complex organisations it may be very difficult to identify the cause, or causes.

The solution is better interpretation of the data, as well as an understanding of variation and 'correlation / causation', for all involved; in particular policy makers.

Ossification - *Organisational paralysis due to an excessively rigid system of measurement.*

There are many examples in the education system which can be viewed as contributing to ossification. Most children stop normal lessons for a period of time every year to prepare for the Key Stage (or optional) tests. Ward *et al* (2002) found that one in seven Y6 teachers spent nearly a year preparing for the KS2 tests. Furthermore, GCSEs themselves severely limit any teaching for several months, and new exams such as the AS levels further impinge on teaching time.

Smith advocates continually reviewing the performance measures; and from the point of view of schools, a regime which used less formal 'high stakes' testing and more diagnostic or formative testing, would probably reduce ossification. Although he warns against too many changes to the system, as this may well make it more difficult to measure changes over time.

Although the various issues have been considered individually, Goddard and Smith (2001) point out that there is a significant degree of inter-dependence. They give the example that setting year

on year targets may help overcome complacency, however, at the same time this may encourage gaming. Furthermore, they discuss the important issue of cost and benefit. Using a small number of indicators will be cheaper and easier to interpret, however this may not give a reasonably accurate picture in terms of performance.

This framework does provide a useful means of judging and assessing PI systems, not only in the public sector, but as has already been seen, in other organisations. The next section looks at three other areas of the public sector; *Health, Rail and Police.* All of these use high stakes indicator systems; and the performance of individual operating units is of public interest and frequently reported in league tables format.

Health

The reporting of performance in health shares many similarities with education. Both are complex systems, and the desired outcomes (eg. healthy / educated people) are not easily or directly measurable; therefore many of the output indicators are used as proxy indicators.

Reporting the performance of the Health service very much represents two extremes. The official data in the form of; *High Level Performance Indicators and Clinical Indicators* includes quite a significant number of PIs, for example for Acute trusts 28 indicators are reported. These cover areas such as clinical effectiveness (eg. Readmission rates) and capacity and capability (eg. Sickness absence rates). The data itself is presented both numerically and in the form of 'league table' graphs, which in contrast to education league tables show confidence intervals, and carry warnings about their limitations; for example in terms of interpreting the data the Department of Health state;

Quality of Indicator – The sensitivity, specificity, repeatability and responsiveness to change of the indicator has yet to be tested

Confidence Intervals - 95%. Some of the values and factors influencing them may be chance occurrences, with values fluctuating at random between organisations and from year to year. Numbers of deaths may be small at individual hospital and Health Authority level. The results should therefore be interpreted with caution and with the aid of confidence intervals. The 95% confidence interval provides a measure of the statistical precision of the rate for an area or institution. It indicates a range which, with 95% confidence, will contain the underlying value of the indicator. If the confidence interval for a hospital's or area's rate is outside the range of the national confidence intervals, the difference between the two rates is considered statistically significant. If the confidence intervals for two rates overlap, in most cases the difference between the rates would not be considered statistically significant.

Effect of case-mix/severity - A number of factors outside the control of hospitals, such as the socio-economic mix of local populations and events prior to hospitalisation, may contribute to the variation shown by the indicators. Differences in case-mix, concurrent illnesses, the complexity of operation(s) and other potential risk factors also contribute to the variation. The data available do not allow

adjustment for any of these factors. This may pose less of a constraint at Health Authority level than at Trust level. We have tried to deal with this constraint by presenting the data in clusters that are similar with respect to institution type.

Other potential confounding factors – The patterns of providing care may vary between NHS hospital Trusts in terms of whether patients are transferred elsewhere for rehabilitation and convalescence before final discharge. Variation between hospitals in average length of stay may lead to variation between hospitals in the proportion of deaths occurring in hospital as opposed to in the community after discharge from hospital.

DOH (2002)

However, at the 'public level' more sensational and statistically dubious figures are produced. For example, the Sunday Times in 2001 produced a league table (based on the official figures and their own survey), with an accompanying article; *Death rates reveal worst hospitals* (Waterhouse 2001). This duly identified the 'best' (UCL), and the 'worst' (Walsall) which by implication was the most dangerous; and significantly there was no 'health warning' attached to this data.

Curiously, Sir Donald Irvine, president of the GMC, was quoted in the article; "This is the most authoritative and accurate account of hospitals and standards that has yet been published". Furthermore, Alan Milburn, the Health Secretary, welcomed the guide stating: "No hospital wants to be bottom of the league. It motivates people. The NHS has acted like a secret society[115]. It has to recognise that people now expect to be treated like consumers".

There have over the years been many claimed dysfunctional effects of the health service indicators. Theorists such as Davies and Lampel (1998) have applied and demonstrated the relevance of Smith's unintended consequences model to the health service. Spiegelhalter (1999) discusses the issue of data manipulation and interpretation, and the evidence that some surgeons are unwilling to operate on high risk patients. In this respect an inquiry (Times 1999) in to the performance of two senior heart surgeons whose mortality rates were significantly higher than the average, found that they (as the senior surgeon) operated (quite appropriately) on the most difficult and 'risky' cases. The danger and potentially dysfunctional consequences being that they might be encouraged to avoid such patients.

In practice the best known dysfunctional effect of NHS KPIs was the 'great waiting list fiddle'. Although many of the 'tricks' were well known, a report from the National Audit Office (2001) provided official confirmation. It was found in many cases that waiting list targets were being met by 'distorting' clinical priorities. For example, treating easy cases in preference to harder complex cases, in particular those that might lead to further complications; 'loosing' or suspending patients

---

[115] Jones (2000) looks at the secret life of the NHS, however she reaches a very different conclusion to Milburn

from the list: changing the classification of operations, and transferring patients to other hospitals. Now waiting times (as advocated by the BMA (1999)) are the key indicator, which although not perfect appear to be less dysfunctional.

In line with Johnson and Kaplan's arguments which were discussed in the last chapter, some of the key indicators may not be suitable for the purpose. For example, survival rates say little, if anything about the eventual outcomes (ie, quality of life). This can be illustrated with reference to a poignant account from a mother of a 'surviving' child at the Bristol Royal Infirmary;

> My daughter Sophie is still classed as a success even though she cannot walk, see, talk, move, has epilepsy, and can't do anything for herself. But because she lived for more than 30 days after the operation she is still counted as a success, and I regard that as a travesty.
>
> (Dobson 1999)

Very much in contrast to education there has been a relatively good level of cooperation between the medical profession and the government in the formulation of PIs. For example, the BMA discussion paper (2000) *Clinical Indicators (League Tables)*, states; "The clinical indicators published for the first time in June 1999 owe their existence to an extensive co-operation exercise between the Health Departments and the Joint Consultants Committee...." (para. 4). Such cooperation and 'ownership' of the indicators by the profession is likely to help moderate potential dysfunctional effects.

In terms of presenting data, besides reporting confidence intervals, increasing interest is being shown in *Statistical process control charts*.[116] They originated in industry, and are being used effectively in a number of medical contexts (eg. Children's mercy hospitals 2002). In the UK Adab *et al* (2002) discuss the advantage of control charts over the current tables;

> NHS performance league tables attempt to portray variation by including 95% confidence intervals.. These make performance league tables visually confusing, and the use of this level of confidence for differentiating between a significant and non-significant difference is rarely appropriate. It also means that that when all providers are offering a similar service in a stable system about 5% (1 in 20) will always be identified as outliers...control charts show no ranking... and the few outliers are easy to identify.
>
> (p. 97)

Railways

On the face of it measuring performance in the rail industry should be reasonably straightforward. Not too surprisingly the most important performance criteria for most users is punctuality (SSRA

---

[116] See Fitz-Gibbon (1996) for a discussion with respect to education

2000). and given the existing electronic infrastructure of track sensors and timetable systems. measuring train times is relatively easy, and unambiguous. However, as has been shown a number of unintended or dysfunctional consequences have been attributed to the performance measurement systems. This issue is considered at two levels; the corporate, and the operational.

During the 1990's British rail was 'broken up' and Railtrack PLC formed to take on responsibility for the infrastructure. Stittle (2002) points out that both theoretically and politically the high degree of deregulation should have provided a more efficient infrastructure. However, as Shaoul (1999) shows it is very difficult to judge how efficient it really was. For example, the subsidies to Railtrack in 1999 were slightly less than those to British Rail for running the entire network. This could be explained by increased investment in improvements to the infrastructure, however this appears not to be the case[117]. Jack (2001) points out that the maintenance costs decreased, and illustrates in great detail how the quality of maintenance also fell. These changes, both in terms of practice and culture, he argues, significantly contributed to the Hatfield crash; "the crash arose from a quagmire of divided responsibility and incompetence" (p. 20).

In essence it would seem that the PIs which measured shareholder return (and therefore directors remuneration) took precedence over those which measured the quality of the infrastructure. Had such qualitative indicators taken precedence, it may be that they would have shown the potentially dangerous state of the track. In effect the key indicators were not appropriate to the complexities of maintaining and improving the infrastructure; furthermore they may have encouraged dysfunctional behaviour, which changed the priorities from safety and quality, to profit. Again the particular context within which Railtrack operated is important, and inextricably linked to this, as Jack (2001) points out; the key people were no longer 'railway men', but rather accountants and casual subcontractors. These changes are very much inline with Hood's predictions for a *New Public Management* (Chapter 6).

In terms of operational indicators the 'passengers charter' indicator system had been shown to encourage dysfunctional behaviour. Perhaps the most infamous 'trick' of how to meet the target was the practice of missing out stations to make up time (BBC 1999). In addition the charter allowed rail operators themselves to declare 'Void' days (SSRA 2000), which would then not contribute to their performance figures. The solution which has been developed by the then Shadow Strategic Rail Authority (now Strategic Rail Authority) was to produce a more comprehensive package of indicators, which are combined to form one overall (proxy) indicator:

---

[117] Monbiot (1999) points out, Railtrack at that time was one of Britain's most profitable companies, and 'paid' its shareholders accordingly

Public Performance Measure (PPM)[118]. In essence it combines figures for punctuality and reliability, for all scheduled passenger services. PPM was introduced in June 2000, and although it looks as if train performance has got worse[119] (Alleyne 2002) it remains to be seen what effect this indicator will have on behaviour as well as performance.

Police

Unlike the railways the primary aim of the police is very complex and subjective, perhaps even more so than health and education. Broad definitions such as lower crime rates and people feeling safer, tend to become rather woolly, and are of limited use in determining priorities. And as Collier (2001) points out this makes it very difficult to measure performance in a valid and reliable way.

There are two main sources of published performance data, firstly at the operational level, the *National Crime Figures* (Home Office 2001). This substantial document brings together data from individual police forces and the National Crime Survey (NCS). The latter helps build a truer picture by aiming to include crimes which might not be reported to the police. The report details the performance of police authorities and command units, in terms of different categories of crime. Whilst doubtlessly useful for internal benchmarking, it also allows commentators to pick and present individual statistics, which whilst being technically correct, may not tell the whole story, for example;

> *It's official: Hackney is more dangerous than Soweto* (Flanagan 2002)
>
> *Burglaries up as police focus on street crime* (Starmer-Smith 2002)
>
> *New figures reveal shocking increase in rural crime* (Hoyle 2002)

Although these might have a similar sensationalistic value to the statistics which show the 'worst' school or hospital in England, they also have some value in communicating policy: for example, the deliberate focus on street crime. Therefore, to gain any proper meaning, national crime figures (NCF) data needs to be contextualised in terms of macro policy and local circumstances.

The second main source of performance data is collected as part of the Local and Police Authorities Best Value Plan (ODPM 2002) – ie. Best Value Performance Indicators (BVPIs). Indicators such as; BV127 – Violent crimes per 1000 population (from the police) and BV121 – Fear of crime

---

[118] See www.railways.dft.gov.uk indicators/toc
[119] There are significant methodological problems in making comparisons with the old system, see Opraf (2000)

(from the NCS)[120] are collected. The prime function of Best Value is as a management tool to encourage the most efficient use of resources; in effect the BVPIs are the police's KPIs.

A number of dysfunctional effects associated with the indicators have been observed. Her Majesty's Inspectorate of Constabulary (HMIC 2000) pointed out that to help meet specific crime reduction targets, a number of crimes had been 'wrongly' classified, as well as others which had not been recorded until after they were solved. Furthermore, the process of targeting particular crimes may well have a negative effect of the detection of other types of crime (eg. a focus on urban crime at the 'expense' of rural crime).

A number of indicators breakdown the process in to separate parts, and this may in some situations be dysfunctional. Collier (1998) gives the example of an indicator which measures the performance of answering 999 calls, taking precedence over an indication measure of public satisfaction with the resulting police action. The fact that the former is easy to record objectively, may be an attraction to those who decide which indicators matter. Overall, Collier emphasises the need for a balanced approach to the use of indicators, with the police service being;

> A values-based learning paradigm which does not abandon efficiency measures altogether but balances the rationality of the performance culture with the natural world of day to day policing... [and].... Such a paradigm is based on values of public service, integrity and justice that are already found within the police service but which may, in the present control dominated regime, become increasingly subservient to quantitative measurement.
>
> (p. 38)

As part of the Treasury's public service productivity review, Clare Spottiswoode (2000) looked at how police performance measures could be improved. She found that in spite of the vast amount of available data, there was a lack of good quality data for decision making. Furthermore, the Best Value system (in particular across the board efficiency targets) encouraged dysfunctional behaviour;

> [There] is a serious gap in the *Best Value* strategy. This gap hits police authorities and forces the hardest because:
>
> - they lack good measures to fulfil their *Best Value* obligations comprehensively to compare performance (including efficiency) with others.
>
> - police authorities and forces themselves do not always know what the scope for efficiency gains is or even where they should be looking for them.
>
> - efficiency targets should take into account the actual position of each force, and thus the gains made in recent years. The Government cannot do this if it lacks a systematic efficiency measure. So, to date, the Government has set uniform, across the-board efficiency targets.

---

[120] A number of cross sector indicators are also included: eg. BV44 – The number of excluded children

(p. 4)

Spottiswoode provides a useful discussion on the difficulties and challenges as well as the limitations, of measuring performance: and this would seem applicable to many organisations and situations, including education:

> There is no flawless method for unambiguously measuring relative police efficiency. All methods necessarily rely on simplifying assumptions and approximations to be able to model the complexity of what happens in the real world. All techniques are therefore open to criticism that they are too simplistic and fail accurately to capture all real world nuances. Such criticisms - although undoubtedly correct in their own way should not rule out the use of techniques that provide insights into, and estimates of, police efficiency. The task is to find the best possible approach to measuring relative police efficiency while recognizing the limitations that any techniques have.
>
> (p. 20)

She rejects simplistic summative efficiency ratios, as currently used, in favour of regression based methods (Stochastic Frontier Analysis and Data Envelopment Analysis)[121]. However, as she points out differential weightings would have to be applied to the outcomes, and given that the original inputs (formula funding) come from regression based models this may in itself be problematical. Furthermore, the setting of the relative weighting would by definition (and appropriately) be essentially political judgments; and as such will subject to political pressure, which may in itself encourage dysfunctional behaviour.

Conclusion

This chapter started off by looking at the theoretical work of Peter Smith and his nine potential unintended consequences of performance measurement systems. This appears to be highly pertinent to education, with most, if not all, of the points highlighting potentially dysfunctional behaviour. A number of issues from this will be used in the detailed survey questions.

The experience of the health sector provides a number of useful lessons for education. The difficulties of measuring the 'right' things are common to both sectors, and much can be gained for looking at how other sectors deal with the technicalities. However, perhaps the most important point is the value of the co-operation between the professionals and government, in terms of the PI systems. This should increase the credibility and perceived ownership of the indicators, and therefore allow improvements and developments, to meet new and changing circumstances.

Rail provides an interesting insight in to the problems of mixing operational and corporate PIs. It is the government's intention to increase private sector involvement in education. Current examples

---

[121] See Spotiswood 2000b)

include for profit companies running LEAs, which will increasingly lead to educational outcome measures being tied to measures such as profit and ROI. For Railtrack this appears to have been a cause in the decline of the quality of the infrastructure. The equivalent in schools would be choices between: new books, shareholder dividends, fixing roofs, directors remuneration etc. Not impossible decisions, but ones which need to made with the benefit of effective and valid performance information[122]. A further issue was NPM, and the consequential casualisation of the workforce, which appears to have led to a reduction in the skills base, and a less committed attitude.

The police service provides a good example of an organisation which can probably never entirely satisfactorily define its key aims, or at least in the sort detail which can provide the necessary 'steer' and direction. Therefore, much of the direction for its operational decisions has to be politically based, and it follows that a robust and trusted PI system is vital for all (stakeholders) concerned. The possible developments (by the treasury) of police performance reporting (Spottiswoode 2000) is in stark contrast to those proposed for education[123] (Mayo 2000). It may well be that presenting KPIs in more sophisticated regression based models, will have less dysfunctional effect than using raw league tables. However, the concept of showing relative efficiency based on financial inputs, although attractive to the treasury, may in reality be of limited value.

A common theme which runs throughout these three organisations is the importance of KPIs and the associated pressure to meet various targets; in effect KPIs and targets have come to dominate much of the thinking, and consequential decision making processes. To develop an understanding of these issues it would seem necessary to look at the wider organisational context, and this takes us back 'back' to the field of organisational theory. Many claims have been made (including by the government) of the value and indeed the need to develop 'learning organisations'. Such structures could well be the key to countering the potential dysfunctional effects of KPIs in the public sector; however, as previously argued in chapter 5 the trend seems to be the other way, towards ever increasing 'scientific' control.

---

[122] The difficulties of the Edison corporation in the US illustrates some of the problems: see Woodward 2002a
[123] Mayo essentially supports the existing system, although he does advocate an 'efficiency' (results costs) measure

Chapter 8

Research design

The reasonable man adapts himself to the world; the unreasonable one persists in trying to adapt the world to himself. Therefore all progress depends on the unreasonable man (Shaw)

Introduction

Returning to the key questions considered by this thesis: *what effect do KPI systems have on the management and organisation of schools and in particular do they encourage dysfunctional behaviour*; the preceding theoretical chapters strongly supported the notion that they have a number of potential dysfunctional effects. Leading on from this position the research element of the thesis has three main aims:

- Identify examples of dysfunctional behaviour
- Give some indication of their magnitude and significance
- Relate the findings to the broader theoretical and organisational contexts

This chapter aims to summarise some of the outcomes from discussions, supervision meetings and other research training which led to the final design.

Research subjects

Initially in 1997 it was intended to just look at English Secondary schools. This was mainly because of the widespread reported dissatisfaction with league tables by many of these schools, and the relative ease of access in terms of research. Shortly after this, league tables for primary schools were introduced, and it was then felt that including this phase, might provide further useful perspectives. Towards 1998 / 9 when the research was being planned there was increasing commentary in the English media on the differences with Scottish schools, both in terms of organisation (eg. Baker 1999 *Back to the future over in Scotland*) and performance (Telegraph 1999 *Scottish schools perform better than the English*); this and other discussions prompted a consideration of how Scottish schools could be included. A brief review of news articles in Scotland (eg. McBain and Louden 1998 and Munro 1999) and discussions with colleagues confirmed that many of the issues surrounding KPI systems were common to both countries.

Raffe *et al* (1999) discuss in some detail the value of using comparisons between the various countries that make up the UK. However, they point out that research too often falls in to the trap of 'lumping' together, Scotland, Wales and Northern Ireland, with England; and that this may lead to findings that are claimed to be generally applicable to the UK, being in reality only relevant to England. They point out that although the different national systems are interdependent (ie. due to UK economic policy), the similarities are more important than the differences; in that the all have a distinct 'Britishness'. For example, the basic structure of the system, many of its functional characteristics and in broad terms the examination systems. On the other hand there are significant differences, for example, in terms of the curriculum, inspection regimes and of course the testing

and KPI systems; and all of these can allow useful comparisons to be made. Additionally, they add that these differences, which perhaps individually may appear insignificant, do tend to be cumulative, in that this may result in larger organisational differences. Furthermore, it may well be that their interaction with other organisational factors further contribute to the differing educational contexts. An example of this is that Scottish primary teachers are largely trusted to administer KPI tests (National 5-14 at the end of primary), where as their English counterparts (KS2) are not[124], and this would seem likely to have other ramifications throughout the two systems.

Using a football analogy Raffe *et al* (1999) make the case for giving 'home international' comparisons a more prominent role in educational research. In terms of policy, the increasing divergence, with for example, the rejection of league tables in Northern Ireland (DENI 2001), and the ending of KS1 tests in Wales (Elliott 2001) would seem to further strengthen the argument and value of such comparisons for the issues considered here. Specifically, for this research the most significant difference is that English primary schools have league tables whereas their Scottish counterparts do not. And the most important similarity is that secondary schools in both countries have broadly similar KPI systems, including public league tables. These differences and similarities are used as the basis for the analysis of the results in chapter 12.

A Qualitative or Quantitative approach?

Much consideration was given as to whether a qualitative or quantitative approach should be taken for the research; both would seem to have their attractions and advantages. Miles and Huberman's (1984) definition that qualitative data corresponds to words rather than figures, initially made this an attractive approach. This was largely because of the inherent complexities and multifaceted issues of KPI systems, which made it difficult to see how the key issues could be 'reduced' to numerical measures. On the other hand to try and 'explain' such complex issues, might in itself create confusion, and the findings might be very much open to subjective interpretation. It will be apparent from the previous chapters that the author's views are somewhat sceptical towards many aspects of the current KPI systems, and this has to be recognised as a form of potential bias.

With respect to a quantitative approach the validity of the findings and any conclusions, are very much dependent on the quality of the data. Returning to Fitz-Gibbon (1996) in chapter 2, data quality can be viewed as a hierarchy, ranging from experiments based on randomised controlled trials (RCTs) down to raw data. RCTs in their simplest form involve one randomly assigned group receiving the treatment or intervention, whilst the other group does not. Therefore, given

---

[124] Currently there are proposals to seal KS tests at the end of the exam to stop teachers tampering with them.

sufficiently large numbers any differences observed will be due to the particular intervention, and not other differences which might exist between individual members of the groups.

There are nevertheless still threats to the validity which are largely due to shortcomings in the design. As previously mentioned the *Hawthorne Effect*[125] may confound particular findings due to the experimental group being treated differently. For example, Gorard (2002) discusses this problem with reference to a piece of research on new methods of teaching maths, however any conclusions were unsafe, because the treatment group were taught by the 'best' teacher who was trying to prove his method. In a similar vein he discusses the short comings of the Hay / McBer research into teacher effectiveness, which again lacked any proper controls. Furthermore, relevant to issues considered by this thesis is the claims that specialist schools are better; is this because of the brilliance of the concept? or perhaps if the intervention is looked at more closely, along with the evidence of improvement (taking account of prior attainment and other contextual factors) the findings might not be so safe.

In spite of the value of well designed experiments the education world has been relatively slow to adopt the principles of RCTs. In the medical world on the other hand, they are far more widely accepted; indeed, Doll (1998) argues that they have transformed medicine more than any other single medical breakthrough. Nevertheless, they are a relatively recent technique, and their widespread acceptance, was neither quick nor easy. Chalmers (1998) and Doll (1998) (and many others) identify a study on the treatment of tuberculosis in 1948 as being the watershed for the adoption and subsequent more widespread use of RCTs. Prior to this, evidence of the best treatments was based on occasional very limited trials and more often the hunches and best judgements of eminent professors of medicine.

This is however not to say that before 1948 some individuals did not try or use randomised trials. Doll (1998) attributes the 'first' attempt to a medical chemist, van Helmot, in 1662. He challenged the academics of the day to compare their treatments based on their theories, to his, based on experience. He proposed that between 200 and 500 sick people should be randomly assigned to the two groups, and that the experiment should be judged on the different survival rates; he even offered a wager of 300 florins. Unfortunately, there were no takers, and it was over 300 years before RCTs became widely accepted, and in turn have been able to show the fallacy (to say the least) of many treatments during the intervening period[126].

---

[125] As discussed in chapter 4 the specific findings of Hawthorne are still subject to debate
[126] As Doll points out even farmers were using RCTs in the 1920s

In the social sciences. Hakim (2000) points out that quantitative experimental research has never been the "dominant mode in empirical social research" (p. 131) in spite of her view that they are of great value:

> [Experimental research] can provide more definitive answers to questions about causal links than do other types of study, and is hence essential for the development of soundly based explanations of social events, behavior and attitudes.
>
> (p. 127)

However, she does make the point that this approach has to take a relatively narrow view which may be an important weakness; "Experimental social research is ... relatively narrow, or focused in the type of information it produces" (p. 127). Clearly, an experiment can only measure what it is designed to measure, and to try and read much more in to it might well invalidate all the findings. Therefore, it may well be necessary to carry out many individual experiments to look at specific aspects of the broader issue. A good example of this is given by Fitz-Gibbon (1996) of a Japanese engineer who set up hundreds of experiments to test just a few variables in a factory. And as she points out; "Social scientists seem to expect a few dozen experiments to yield definitive conclusions even though they are dealing with far more complex 'production lines." (p. 19).

At the policy level, Slavin (1989) points to the dangers of not endorsing the principles of evidence based educational policies, in that education policy will tend to move from fad to fad. And all too often these are based on political expediency and prejudice rather than any proper evidence. Furthermore, as Fitz-Gibbon (1996) points out unless educational policies are properly evaluated, there are many losers besides the children concerned;

> If policies were routinely piloted in the framework of an experimental design, we might gradually begin to understand the costs and benefits of decisions which affect millions
>
> (p. 18)

In practice at the policy level she illustrates this with the case of nursery education. An issue fraught with many impassioned feelings and views on all sides. And therefore it is essential that the policy is based on good evidence, which with carefully designed experiments is realistically obtainable.

Likewise, Slavin (2002), points out that, when looking at the impact of proposed programmes there are, "few alternatives to well-designed experiments". Nevertheless, he does recognise that in some situations RCTs might not be viable: "In some policy contexts, experiments are impossible, and well designed correlations or descriptive studies may be sufficient" (p. 18). Furthermore, and relevant to many of the issues considered by this thesis, he points out that, "Correlation and

descriptive research is essential in theory building and in suggesting variables worthy of inclusion in experiments." (p. 18).

Gallagher (2002) too points to the advantages of experiments, such as being able to carry out valid replications. He does however have some reservations on what RCTs can achieve in some situations, and points to the importance of context. For example, he argues that even with ideal RCTs, 80% or more of the influences will lie outside the treatment or experimental programme and; "That this is in stark contrast to... medicine, where the sheer potency of the drug can overwhelm any contextual factors in the study." (p. 3). For some medical interventions this distinction may be true; for example, the effect of different drugs on a specific infection. However, there are many very relevant contextual factors in the treatment of *complex* psychiatric disorders, and indeed the challenges of designing effective RCTs would be similar to many educational interventions. And at the policy level there would be very little if any differences between education and health.

Although, there are obviously difficulties in designing effective experiments, some theorists appear to take something of a defeatist view. For example, Levacic and Glatter (2001) when looking at how an Evidence – Informed Policy and Practice (EIPP) model might be used to review educational management research, make the following points;

> Moving further along the continuum there are interventions that are universally applied to an education system (such as local management of schools, national curricular changes, performance management of teachers) so that there are no control groups which are not experiencing the intervention. Unfortunately for EIPP research, much educational intervention has been of this type.
> (p. 20)

Although controls in the true experimental meaning (ie. proper RCTs) may well indeed not be achievable in some educational interventions, there are nevertheless still opportunities for reasonably valid 'comparison groups' to be formed. For example, in terms of the national curriculum, the many private schools which have not adopted it, and for performance management, other training organisations which are similar to schools, but do not use the DfES model. Some interventions will be specific to certain L(EA)s or groups of schools, for example special needs policies, and in such situations the quality of the comparison groups may be close to true randomisation. Furthermore, as is attempted by the research in this thesis, other countries of the UK (or perhaps wider) can be used as comparators[127].

---

[127] Indeed, retrospectively in terms of league tables the recent decision by 19 leading independent schools to opt out (Canovan 2002), creates a very useful comparison group, and one that will be subject to 'market forces'

Although this section has pointed to the advantages (and some of the problems) of experimental research, this is not to suggest that qualitative research is not useful or does not have an important part to play. In some respects qualitative research is more difficult to define, and the quote below from Hakim (2000) helps clarify the issues:

> Qualitative research is concerned with individuals' own accounts of their attitudes, motivations and behaviour. It offers richly descriptive reports of individuals' perceptions, attitudes, beliefs, views and feelings, the meanings and interpretations given to events and things, as well as their behaviour.
>
> (p. 34)

On the face of it as mentioned before there are a number of attractions in this approach to the issues considered by this thesis, in particular the critical importance of the individual respondents feelings and perceptions, which may not easily be measurable in a pre-planned and defined experimental setting. Within social and economic research, Hakim points to the dichotomy between quantitative and qualitative research, and suggests that there may be an incorrect perception that 'serious' research is biased towards quantitative methods, in that this approach is viewed as necessary to be able to 'prove' a point. She argues that this may not be the case, with much depending on the nature of the questions and the particular settings, and suggests that there is an increasing awareness and valuing of rigorous case studies, and points to the arguments that; "...truly experimental research in natural settings is virtually impossible in the social sciences" (p. 13). As argued above the particular intervention (or component parts) and context, as well as the definition of 'truly', are the critical factors; and with careful and creative design it may be possible to overcome these difficulties.

Hakim (2000) also points out that even once quantitative data has been obtained there are many issues surrounding its proper analysis and interpretation; "...there is growing recognition of the limitations of statistical analysis and linear logics when considering change processes that involve multiple causation or multiple outcomes.[128]" (p. 13). On the other hand Meyer (1995) points out that such potential threats to validity can be significantly reduced by using and combining for analysis, a number of both interventionist and non-interventionist studies. An example of this in education being the EPPI[129] reviews, one of which has recently provided some interesting views on the issues considered by this thesis (Harlen and Deakin Crick 2002)[130].

Whilst there are arguments that a qualitative approach would be able to deal with many of the questions raised in the previous parts of the thesis, and indeed to provide some very rich insights of the issues surrounding KPIs, the key problem remains; how to get a 'handle' on the important

---

[128] She mentions complexity and chaos as being potential ways forward
[129] EPPI - The Evidence for Policy and Practice Information and Co-ordinating Centre
[130] Specific reference to the review is made in chapter 14

issues and to assess their significance. And it was difficult to see how this could be achieved with a qualitative approach. Furthermore, in terms of disseminating the findings and hopefully comparing them with other findings, there are advantages in some form of quantitative analysis.

In spite of the advantages of a quantitative approach there are nevertheless a number of specific practical difficulties; for example, the blanket imposition of KPIs and their close relationship and interdependence with many other organisational factors. Whilst problems such as these can be overcome by identifying and 'measuring' various component parts of KPI systems (eg. Target setting, feedback systems, league tables), the one insurmountable problem is that of timing: in effect (stating the obvious) the experiment needed / should have been done before the intervention[131]. To help address these difficulties, Cook and Campbell (1979) advocate a quasi-experimental design; in effect a middle way. This approach can allow causal relationships to be identified and measured without control groups and the random assignment of the intervention: and therefore such an approach would seem relevant to this thesis.

Although the previous theoretical chapters have identified many potential issues or questions worthy of research, there are still gaps remaining in the clarification of what the key issues are and what issues or questions should be 'measured'. Thietart *et al* (2001) usefully suggest that qualitative research is best suited to establishing the key questions and forming hypotheses, and quantitative research is more suited to testing these hypotheses.

Therefore, in the final design this approach of qualitative then quantitative research is followed. An initially mainly qualitative approach is used to supplement the theory to help identify the key issues, and from that form the hypotheses and research questions. Then a mainly quantitative approach is used to test these questions. As the use of RCTs was not possible, a quasi-experimental design, based on a differences and similarities between the Scottish and English education systems was used. This aimed to assess the effects of KPIs and their component parts, in terms of differences between; country, phase and post[132].

---

[131] Although, given the recent removal of the interventions (ie. league tables) which are of the interest to this thesis in Northern Ireland and Wales, it would be possible to do an experiment in reverse.

[132] However, this is not to say that more rigorous RCTs should not therefore be used for other aspects of KPI systems, for example the proposals to impose targets on primary schools, which have not yet been introduced: far from it is highly desirable or essential that this occurs before implementation. It is worth noting that Earl *et al* (2001) who carried the government's review of the NLNS advocate such an approach (see chapter 14)

Possible research instruments

Details are given below of some of the main methods of collecting the qualitative and quantitative data which were considered.

Face to face interviews

Although these could be used to gain quantitative data, for example, interviewer completed questionnaire, this would not be a very efficient or a productive use of interviewer time. Furthermore, most heads (by their very nature) would want to elaborate and discuss the questions. Therefore, such interviews in schools (as opposed to other setting such as random shoppers) would seem more appropriate to gaining qualitative data. Something of a compromise can be achieved by using semi-structured interviews which provide quantitative data, and can also provide qualitative data by allowing interviewees to qualify and contextualise their responses. A further advantage of this approach is that some degree of learning can take place due to the contact with those involved.

There are however a number of potential disadvantages. In terms of gaining qualitative data it would be very expensive in terms of time and costs to be able to interview sufficient schools in a wide enough area to gain reasonably representative data. There would be difficulties maintaining consistency throughout the process, largely because of the 'learning' element; indeed this could occur both ways, with a danger of introducing interviewer bias.

Telephone interviews

A number of the problems of face to face interviews could have been addressed with telephone interviews. It is possible to cover a wide geographic area, the reliability could be improved (and bias reduced) with the use of a standard script. In addition, more 'free ranging' qualitative questions / responses could also be allowed. The actual cost per response would be relatively low and the collection period short (Frey 1989).

However, given the nature of some of the questions and potential sensitivity, the lack of anonymity may discourage full and honest answers. Furthermore, the subjects would have to be contacted at work and this might not be the best time / place to give full and considered answers, and some questions might require access to other information sources.

Longitudinal case studies

These could be carried out with a number of schools, and would allow a more complete picture to be gained, in particular at a time of organisational change. In the case of English primary schools this could have been started before the introduction of league tables, and carried on for the first few years of their existence. Furthermore, first hand observations could have been more easily made on

the behavioural effects of KPIs, rather then having to rely on observations from heads and teachers. However, there are a number of potential disadvantages in this approach, in particular the very limited number of cases that could be studied, as well as the danger of going 'native' and loosing some degree of objectivity.

Focus groups

This approach was given some consideration, and although potentially a very useful way of exploring some of the complex issues, its use for this research would seem to be rather limited. At the early stages of identifying the key questions it could have produced a rich (and probably colourful) supply of qualitative data. However, for the main part of the research there would be a number of difficulties. By definition most if not all of the participants would know each other, and indeed may well have previously discussed some of the issues, which might inhibit frank and honest answers. Furthermore, to be of any use in terms of comparative research (ie. England and Scotland) participants from the two countries would have to be brought together, and this would create logistical problems. In additions, there are a number of significant differences in terminology and practice which would probably have caused some difficulties. Telephone conferences could potentially overcome some of the issues such as being able to link people from different areas, however co-ordinating six heads, for example, to be available at the same time would be virtually impossible (eg. teaching commitments for primary heads).

Email / internet conference

This could take the form of distributing quantitative questionnaires or enabling qualitative discussions. Shortly before this research, Visscher *et al* (1999) carried out some very interesting and relevant research[133] by using an email conference between various experts in the field. In terms of obtaining quantitative data there would not seem to be any significant advantages over more traditional postal surveys; indeed there would be disadvantages in terms of controlling the distribution, furthermore, some heads may not have access to suitable facilities. As previously mentioned an approach was made to the NCSL to use their *Talking Heads* internet based conference system which potentially could have been very useful and almost certainly would have provided interesting insights. However, in spite of initial enthusiasm with a claimed desire to support the 'research community' this was not allowed to proceed (they declined to give any reasons).

Postal questionnaires

This approach is of course a very common method of acquiring quantitative data. It has a number of advantages in that a large number and wide range of respondents can be contacted throughout

---

[133] *Evidence on the intended and unintended effects of publishing school performance indicators*

the sample area, at a relatively low cost. The questions can be made consistent both in content and context, and the responses can, if necessary, be made anonymously. The responses can also be easily coded for computer analysis, and there is far less potential researcher bias. In addition respondents can complete the questionnaires at their leisure, and do not need access to computers or email.

On the other hand the questions will tend to be more 'closed' and there is less opportunity to contextualise responses. This can reduce the chance of issues being raised which had not been previously considered. Response rates tend to be lower than some other methods such as telephone interviews, and consequently there are difficulties in dealing with non-returns.

Final design

From the above discussion it can be seen that there was not a clear simple, one best way to approach the research; all the instruments having pros and cons. Therefore a number of these different approaches were used, with the aim of combining their respective strengths which would give the most effective overall design. To achieve this the research was divided in to three stages: *Preliminary, Main and Follow-up.* These stage are summarised below, and then expanded in the chapters which follow.

Preliminary

Given my lack of detailed knowledge of the key issues facing schools and the lack of specific previous research in this area, the preliminary research was used both as a learning exercise, and for testing some of the potential questions which had arisen from the theoretical aspects of the thesis. Face to face or telephone interviews were used with a number of heads from English and Scottish schools, and a semi-structured format was used to help maintain consistency.

Main research

It was decided that this should be predominately a quantitative exercise because by this stage quite a number of key issues had been identified which could then be tested. A postal questionnaire was used to gain quantitative data, in addition some degree of qualitative responses could also be made on the form.

Follow-up

At the time of the main research there were many changes occurring in the education systems, both in England and Scotland, and there was a concern that these may have adversely affected the responses. Therefore, a follow-up telephone survey was carried out with a number of the respondents (to the main questionnaire), who were re-asked some of the key questions. In addition

some more qualitative discussions took place which allowed some degree of 'longitudinal' assessment to be made.

Chapter 9

Preliminary Research

If everybody is thinking alike, then somebody isn't thinking. (General Patton)

Introduction

The preliminary research consisted of face to face and telephone interviews, carried out in 1998, and it had four main aims;

- Identify the key and most significant issues facing schools in terms of their performance indicator systems
- Extend and develop my own understanding of the use of indicator systems in English and Scottish schools, in particular their place in the broader management and organisational context
- Try out some of the potential questions and answer formats
- Identify other significant issues related to PIs and their broader context which may not have been identified or felt to be of importance in the previous theoretical section

For the English part, four nearby (North Yorkshire LEA) schools were contacted and agreed to take part (Secondary *S1* and *S2*, Primary *P1* and *P2*)[134]. Based on reputation one (*S1*) was considered to be very successful, another (*P1*) was in a relatively difficult urban area, and the other two were 'typical' schools (for the LEA). However, it should be noted these schools were not nationally representative, indeed the LEA is far from 'average' or particularly typical. Two Scottish schools also agreed to take part (a secondary *S3* and a primary *P3*). Both were from a large town in the Dumfries and Galloway EA area, and both were considered successful.

For the English schools the research was carried out by face to face interviews with the head (some follow up contact was maintained). For the Scottish schools telephone interviews were used. These took place on a number of occasions over a period of several months, indeed these two kind and generous heads were literally operating an 'on-demand' telephone consultancy service. Much of this was necessary due to my lack of familiarity with the Scottish education system. In addition, contact at this stage was also made with the Education Authority (Dumfries), HM Inspector of schools (HMI), Scottish Qualifications Authority (SQA) and the Educational Institute of Scotland. Considerable help and resources were kindly provided by these bodies. Together with the school contacts it was possible to build a useful picture of the issues surrounding KPIs, as well as other broader issues relating to the Scottish education system.

Prior to the visit (or telephone call) the schools were sent a brief outline about the research as well as details of this preliminary exercise (*Appendix 1*). In addition for each school reference was made to the published KPIs and Ofsted / HMI reports. A semi-structured format was used in the form of an interviewer completed quantitative questionnaire. The preliminary research questionnaire is

---

[134] These school identifiers (eg. S1, P3) are used in this chapter to attribute the different quotes or perceptions

shown in *Appendix 2*, and appropriate adjustments were made to the terminology for the Scottish schools.

The questionnaires aimed to build up a picture of the schools, firstly in general terms, then more specifically with reference to their own indicator systems. Following this external indicator systems were considered including KPIs, and finally a number of questions were asked about the effects of indicator systems. Although specific questions were used, this was primarily to structure the interview. Where appropriate these were expanded and other points and issues were discussed; in effect the interviews were quite loosely structured.

Findings

The process was extremely valuable in terms of the four main aims outlined above, and although quantitative values were assigned to the questions, given the size and nature of the sample, and the purpose of the exercise, no valid conclusions can, or were intended, to be drawn.

Two of the schools (*S1, P2*) used Durham CEM systems and were positive about the benefits. All the Heads easily (and enthusiastically) identified and discussed in qualitative terms their school's main aims and key success factors. Indeed, for most such an exercise was almost second nature. Two schools did not consider the use of specific internal indicators for day to day management important (*P2, P3*) with these heads preferring to rely on their judgement. However, more use was made of such indicators for longer term (strategic) management.

All of the Heads stated that the collection and compilation of indicators required a significant amount of work, however there were differences in their responses as to how the indicators were used by others, such as staff, governors and parents. Two (*S2, P1*) said that their staff were all aware and interested in the indicators. In terms of the governors, one head (*S1*) said they were only interested in the KPIs and felt that any other performance information was the responsibility of the head.

There were varied opinions in terms of the use of indicators in the planning process. In general there was a belief that they had some part to play but this varied from school to school. A similar finding occurred for the notion that league tables are of benefit and help raise standards, with much appearing to depend on how the particular school performed. All of the heads knew of 'unfair' cases, and saw limitations and dangers of using free school meals as a basis of comparison.

There was some disagreement as to whether external indicators provided a valid assessment of schools and education in general. To an extent the individual school's relative performance in the league tables, and Ofsted experience may have been a factor. There was a feeling that external indicators were broadly understood by other stakeholders. Interestingly, there was some suggestion (*S1*) that the local public perception could be influenced by how the local media interpreted and reported the data, (mainly league tables).

The question as to whether KPIs helped improve education in general, was too simplistic on its own. It raised many issues, such as being useful as a management / incentive device, with much depending on how the particular school perceived their results. There was a view that they could contribute to strategic planning, but there was a problem of 'shifting goal posts' due to frequent changes in government policy. There was also a general feeling that the basic principle of increased public accountability did help raise standards, and that an indication of value added would further improve the figures. There was however some confusion (*P1*) on the technicalities of value added calculations[135].

Most heads did find that the external and internal systems gave a broadly similar picture, although again how the school performed externally appeared to influence their perception. Generally, it was felt that governors, staff and parents saw a 'similar' picture. There was not a clear view as to whether external indicators tended to 'override' internal indicator systems, although the heads in general felt they did not 'override' their professional judgements. Furthermore, it was generally felt that schools could and should still pursue objectives which were not necessarily measured.

Whilst it was felt that the indicators did give some indication of current performance, it was not felt that they gave a significant indication of future performance; an interesting view, given the importance many parents appear to place on league tables, when choosing schools. There was a varied response as to whether schools tended to concentrate on 'borderline' children. Overall, there was support for national target setting, although a more varied response as to how much local negotiation should be allowed.

All of the heads said that their indicator systems did not encourage competition between teachers, although one commented that he wished it would (*S1*). There was a varied response to the subject of performance related pay, as well as the idea of linking school performance to the level of funds it receives. There were several comments about some of the official information which they

---

[135] In particular whether this was done on a whole school or individual pupil basis

received being out of date (*P1*). For example, it was suggested that much of the new Panda data was already known to schools, being largely based on their own data and old Ofsted reports.

Several of the Heads (*P2, P3, S3*) wanted more information on the progress of individual pupils, and other 'softer' data, for example parental attitudes. The two heads (*S1, P1*) who were using third party systems (Durham CEM) found these to be useful in this respect. Finally, most of the heads did not feel that their indicator systems were properly used by the (L)EA as part of their decision making and strategic planning process.

Conclusion and Hypotheses

The exercise was very useful in terms of relating theory, and in particular government policy to the actual day to day practice of managing schools; as well as helping to build an understanding of some the key issues facing heads. For example, the difficulties faced by the school (*P1*) which had been identified as having weaknesses in being able to communicate to the LEA their strengths and effectiveness (at the individual pupils level) in spite of their difficult circumstances.

In general the questions which were asked were found to be relevant to the aims of the thesis, furthermore a number of the other issues raised were subsequently used in the main research questionnaire. One particular problem which arose, was the use of a seven point scale, (from Strongly Yes to Strongly No), with the respondents appearing to spend too long weighing the most appropriate level of response; furthermore one person's 2 would be another's 3 and so on. This issue was addressed in the main survey by using five point scales.

One quite surprising finding was the relatively high degree of interest that the heads had in indicator systems *per se*, and in general they found them to be useful in terms of the organisation and management of their schools. Therefore, it was concluded that the main questionnaire which was to follow should contain a range of questions which covered both the use of indicators as well as their potential dysfunctional effects.

Main Hypotheses

At this stage six main hypotheses were established, and these were used as the basis for the main questionnaire:

H1 - KPI systems overall encourage dysfunctional behaviour: The majority of the theory considered in the first part of the thesis, and the experiences of schools and other organisations suggest that the KPI systems will have dysfunctional effects on schools.

H2 - There will be overall support for good performance information systems: Schools made it clear that they wanted to have access to quantitative performance data, and to use it for their organisation and management.

H3 - Secondary and primary schools have similar views on the use of performance data for school management, but secondaries find KPIs more dysfunctional: Secondary schools tend to be higher profile and are more in the public and media eye. This will encourage them to engage in more dysfunctional behaviour to meet their targets. Furthermore, there is less scope for such behaviour in primary schools.

H4 - English and Scottish schools have similar views on the use of performance data for school management, but English schools find KPIs more dysfunctional: The higher profile and more widespread publication of league tables, suggests that the English schools will be under more pressure, and this will lead to increased dysfunctional behaviour.

H5 - Teachers and heads have similar views on the use of performance data for school management, but teachers find KPIs more dysfunctional: The role of heads has significantly changed, in part due to the reforms in education as well as the introduction of KPI systems. This has resulted in pressure on heads to take a more 'managerialist' role of directing and pressuring teachers to meet KPI targets.

H6 - Schools which feel they perform well will attribute less dysfunctional behaviour to KPI systems: It is expected that schools which do not perform so well will be under greater pressure to meet KPI targets and therefore more likely to find the indicators dysfunctional.

Chapter 10

Methodology for the main survey

Mission Statement A talisman hung in public places to ward off evil spirits (Shapiro)

Introduction

Chapter 8 suggested that postal questionnaires would be the most suitable instrument to test the issues and hypotheses which had been previously identified in the theoretical section and preliminary research. Although this would give predominately quantitative results, it was felt that there would also be an opportunity to gain some qualitative evidence. It was decided that these questionnaires would to be sent to a number of randomly selected schools in England and Scotland. in addition extra copies would be included for a randomly selected teacher and the chair of governors (English Schools only) to complete. This was aimed to help assess the hypothesis that teachers find KPIs systems more dysfunctional then heads, and overall to provide some degree of triangulation of the data.

In effect this approach relied on the views of those people closely connected with the school. There is an argument that they may not be in the best position to make judgements about the effects of KPIs on their schools. For example, target setting and league tables have clearly proved unpopular for many schools, and therefore the respondents may, to an extent, be biased against such systems. This potential bias will not be such a problem when groups, for example primary and secondary schools are compared, but should be borne in mind for results which show the overall responses.

On the other hand there is little doubt that these people are in a unique position to assess accurately and honestly what actually happens in their schools and classrooms. Furthermore, there is to a degree, a feedback loop between the views and feelings of the staff, and what actually occurs in the schools. Therefore, it was felt that heads and teachers are the best judges of how KPI systems influence the behaviour and organisation of schools, although this is not to say they other perspectives, for example, advisors and inspectors would not also have been relevant.

The questionnaire

The detailed questions were constructed with reference to the main hypotheses and the many specific issues which had been identified in the theoretical section. A draft questionnaire was produced and four head teachers in England who had not taken part in the preliminary research were contacted and agreed to trial it and make comments. *Appendix 3* shows the comment sheets. The overall feedback was helpful and positive, and from this the main English (primary and secondary) questionnaire (*Appendix 5*) was finalised. This was subsequently adapted for Scottish schools with the help of the two heads with whom contact had been maintained by making specific

changes to the terminology (eg. Standard grades rather than GCSEs) and in some cases removing questions (eg. references to league tables for the primary schools).

The questionnaires were sent to the heads with a covering letter (*Appendix 4*) and a return envelope. They were asked to complete one themselves, and to pass another to a randomly selected teacher. For the English schools this was asked to be the teacher with a surname beginning with M. (or the next letter after this in the alphabet) and with at least 2 years teaching experience, and in Scotland the letter G was used. In addition English heads were asked to pass a copy to the Chair of governors.

Question formats

Three different formats were used:

Likert Scale – The seven point scales which had been used in the preliminary research were replaced with five point scales, because it was felt that they slowed down the process, and gave very little, if any, extra precision. The possible responses were: YES definitely; YES mainly; Neither or Yes & No; NO mainly; NO definitely. The results were coded using a 1 – 5 scale. Four and six point scales were also considered as a means of 'forcing' a one way or the other response. However, it was quite conceivable that there would be 'genuine' middle responses, and it was felt that to remove this option might have alienated some of the respondents, who were of course very experienced at completing surveys. In the event the distribution of the responses (chapter 11) shows very little evidence of 'middling through'.

Semantic differentials – These consisted of a general statement or question (item), with a number of pairs of responses separated by 10cm lines. A cross was placed at the appropriate position on each line to signify the respondent's feelings. This was originally coded on to a 1 to 10 scale. However, after trying a number of different analyses it was found that virtually identical results were obtained when the scales were reduced to 5 points. Given the advantages of being able to make easier comparisons with the other formats all of the semantic differential scores are reported in this thesis on five point scales. The items were arranged with the potentially dysfunctional responses alternately on the left and right to discourage a neat line of possibly 'self-fulfilling' crosses.

Point score – These were simply scored on 1 to 5 scales.

For all of the questions a separate response of Don't Know (d/k) was allowed, because it was quite conceivable or indeed probable that some of the respondents would either not know the answers or for personal reasons prefer not to answer them. This was considered preferable to having respondents simply ticking the middle box, which would have reduced the accuracy of some of the statistical analyses.

Respondents were also invited to make further comments, both whilst completing the questionnaire, and in a separate section at the end. The returned forms were coded and entered into Microsoft Excel for the initial data checking and manipulation. This was then transferred to SPSS for the analysis. In some cases the scales were reversed to allow easier and more obvious comparisons to be made. The data from the Likert scales and point score should arguably be treated differently to the Semantic differentials, ie. discrete and continuous. However, given the nature and format of the questions it would seem most likely that the respondents would answer them all in a similar way, therefore in terms of the analyses all the questions are treated in the same way[136].

Distribution

For the English schools a randomised list of schools was produced by the CEM centre in Durham. From this one hundred questionnaire packs were sent to secondary and primary schools. For the Scottish schools I was advised that it would be unlikely that many would respond directly to questionnaires, and that it was best to get prior agreement with their EAs in advance.

Therefore, four EAs were selected on the basis of being reasonably representative of Scotland as a whole, in terms of their urban / rural mix, and their average exam performance (*Table 7*). In terms of performance at Standard grades, together all four EAs achieved 'average' results in 1998, which would suggest that the primary schools in the four areas would also be 'average'.

With the EAs agreement a total of 76 primary and 58 secondary questionnaire packs were sent out to randomly selected schools in the four authorities. The randomisation and addressing was achieved by using a spreadsheet file (CSV) of all Scottish schools and their EAs which was provided by the Scottish office.

---

[136] See Cramer 1994 p 53 for a discussion of this point

Table 7: % S4 Gaining 5+ Standard grades at 1 – 2

EDUCATION AUTHOROTY	RESULTS
Dumfries and Galloway	35
North Ayrshire	28
South Lanarkshire	31
West Lothian	27
Mean of all four authorities	30
National Mean	30

Source: *Examination Results in Scottish Schools 1996 – 98.* The Scottish Office

Responses

Questionnaires were returned directly to a freepost address at Durham University and the various response levels are set out below.

Table 8: Questionnaire return pattern

COUNTRY	PHASE	SENT	RETURNS			
			% Schools	Heads	Teachers	Governors
England	**Secondary**	100	34	34	19	8
	Primary	100	28	28	11	8
Scotland	**Secondary**	58	30	17	9	n/a
	Primary	76	28	21	7	n/a

Notes: Several of the primary schools just had a Head / Teacher

The overall response rate from schools (heads) of about 30% was a little disappointing, and the reasons for non-returns are important. Fourteen schools overall made 'apologies' and wished the research well. The most frequently cited reason was pressure of work, several (2) were preparing for Ofsted inspections and several (5) said they were overwhelmed with requests and would only respond to 'official' surveys, such as those from their education authority or the government. Of the responses five preferred to be anonymous[137], although none of their answers were in anyway extreme. The response rates from the teachers and governors was also disappointing, but it is difficult to know in how many cases the questionnaires were actually passed on. In the follow up

[137] This was achieved by removing the reference number on the form, also about 2% preferred not to explicitly name their school

one head commented that he felt that he was the only one able to make a proper response to the questions posed!

Consideration was given to follow-up letters. However this was rejected as these may have gone to the anonymous respondents and it might have unnecessarily alienated others. It would still of course be interesting and important to know the reasons for the non-responses. That said, the quality of the responses was very high, with clearly the vast majority, if not all making carefully considered responses, and in many cases additional useful 'free text' comments were made on the forms. About 35% added comments (free text) to the additional information box at the end of the form, 40% offered to take part in the follow up research, and 50% requested further information by post/email.

To help put these responses into a broader context, a survey (carried out by NFER on behalf of the government) on how schools use National Curriculum Test Results (Ashby and Sainsbury 2001) yielded a 53% response rates from either heads or teachers with specific responsibility for the tests. The survey by OISEUT[138] for the DfES (Earl *et al* 2001) on the National Literacy and Numeracy strategies yielded a 49% response rate by heads and a 20% rate for teachers.[139] A survey by NUT on Education Action Zones yielded a 13% response rate from school NUT representatives (Theakston *et al* 2001).

Representativeness of sample

It was important that the sample was reasonably representative, ie. not predominantly 'good' or 'bad' schools, or those that feel the tables are very unfair to them. To assess this two introductory questions were used; firstly the respondents were asked how their school performed, compared to other schools in their authority area, and secondly respondents were asked how fairly this reflected their schools overall performance. The responses to the first question were not expected to be an exact representation of the performance, but more importantly to indicate how the respondents perceived their school's performance. In a similar way the second question was designed to indicate the respondent's feelings, towards the fairness of their position.

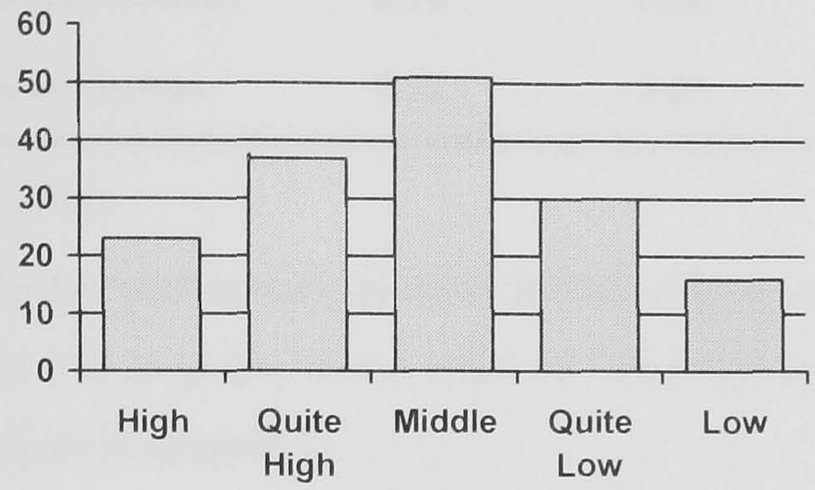
Figure 8 shows that the majority of the respondents considered that their schools were around the middle in terms of relative performance, with far fewer at the extremes. In reality the graph 'should' have been flatter, a check on a number of the responses against the results in the performance tables suggested that the respondents tended to describe themselves as 'average'

---

[138] Ontario Institute for Studies in Education, University of Toronto
[139] The teacher part of the survey was abandoned because of the low rate

rather than very good or very bad; put another way the scale was not interpreted in the proper proportion (20% steps). Nevertheless, the results do suggest that the respondent schools are reasonably representative.

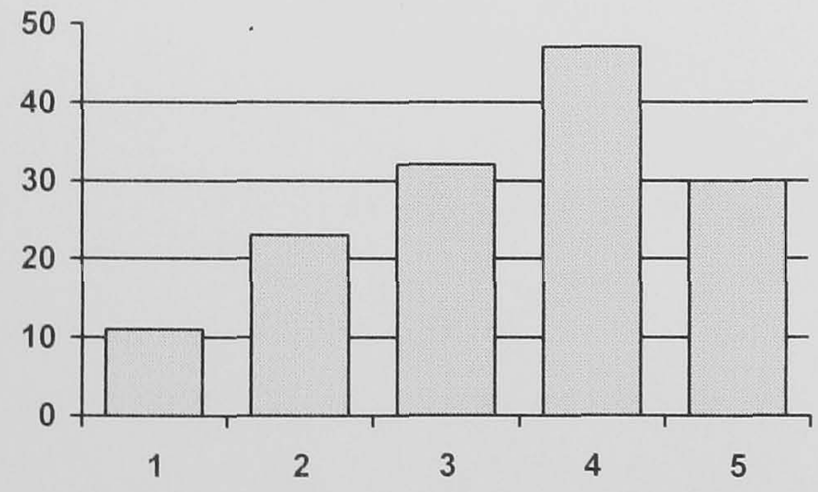
Figure 8: Schools relative performance



N=157 Question: How would you assess your schools position?

Figure 9 shows that a majority of the respondents considered that their KPIs were a fair reflection on the overall performance. This is important as it suggests that few would automatically 'have an axe to grind' due to perceived (or actual) unfairness of KPI systems on their school; therefore it would seem less likely that their responses would be inherently biased.

Figure 9: How fair are KPIs?



N=143
Question: Is this a fair reflection of your school's overall performance?
Score 1 very unfair to 5 very fair

The results of these two questions have been further analysed to show how the results differ between the English and Scottish schools in *Table 9.* The results suggest that there is little difference between the two countries, with respondents from both being relatively 'average' and happy that their results are reasonably fair.

Table 9: Tests of Significance for differences in position and fairness

QUESTION	MEAN		STD. DEVIATION		SIGNIFICANCE	
	England	Scotland	England	Scotland	t Test p value	Effect size
Relative Position	2.79	3.02	1.28	0.97	.26	.20
How fair is this	3.32	3.65	1.29	1.00	.12	.28

N = 157/143

A further preliminary question looked at how long the respondents had been in post. The range was from 1 to 25 years, with a mean of 7.66 years and a median of 6 years, this data is used in the analysis in chapter 13.

Analysis of the results

A total of 162 responses were received which are analysed in three stages over the next three chapters. The first provides an overview of the data with descriptive statistics (in graph form) of all of the responses combined, in addition a number of the comments made by the respondents are also reported. The second looks in more detail at the differences between groups of respondents, based on Country, Phase and Post. Finally, some of the relationships or how the responses correlate are explored, along with details of the follow up telephone research.

Chapter 11

Descriptive analysis

Statistics are like bikinis. What they reveal is suggestive, but what they conceal is vital.

(Levenstein)

Introduction

This chapter analyses the distribution of the results from all of the questionnaires (n=162), with the aim of giving a broad overview of the attitudes towards KPI systems. The layout is based on the main (English) questionnaire *Appendix 7*. As mentioned in chapter 8 a number of changes were made to the Scottish questionnaires because of technical or terminological differences. For example, references to the effect of league tables were omitted for Scottish primary schools, and the Scottish KPI of the percentage of S4 pupils gaining Higher grades in S5 which has no English equivalent is reported separately. Where appropriate the differences are highlighted in this chapter, rather than cross referencing three separate questionnaires in the appendices.

To present the data three graph formats are used for the different question formats: bar charts for the Likert scale questions, area graphs for the semantic differentials and column graphs for the point scores. The mean and number (n) of responses are also shown, although n does not always equal 162 because some questions only applied to some schools, and any respondent either leaving the question, or marking it *'don't know'* has been omitted. All of the responses are on a scale of 1 to 5, therefore 3 is the mid point and 1 would signifying all respondents answering yes definitely, or completely agreeing with the first term in the semantic differentials (5s would be the opposite). Some of the additional comments from the survey forms are also reported, with the aim of illustrating particular points in the discussion. Individual respondents are cited as RS nnn.

Use and awareness of KPIs

The first group of questions aimed to gauge how important KPIs are to the respondents and their schools. *Survey 1* shows not surprisingly that the majority of the respondents felt that they had a good idea of how their own school was performing, although *Survey 5* suggests that only a small number of schools felt that their relative performance with others in the area was important. Several comments were made about the need for schools to co-operate with each other. Key to the government's reform agenda (chapter 7) has been to increase the prominence of KPIs and the level of competition between schools, these results suggest that there may be some resistance (intentional or not). This would appear to confirm Cowie's (2000) findings:

> Headteachers in this study not only rejected the values of the market philosophy with its emphasis on competition and choice, they endorsed cooperation, with duties and responsibilities towards each other on the part of both schools and the authority.
>
> (p. 21)

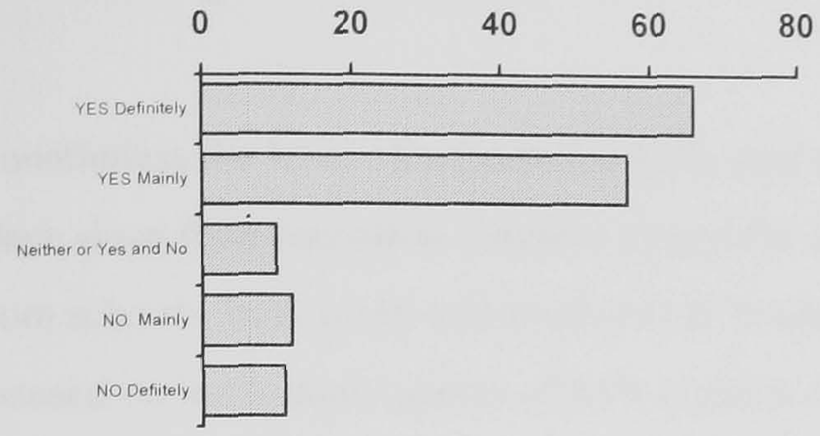
Survey 4 suggests that only a few schools see their main aims in terms of KPIs, and therefore what is measured should not be the most important issues for schools. An important part of the government's drive to raise standards has been to increase accountability to parents, in line with the principles set out in the citizens' charter (chapter 5). The results from *Survey 2* and *Survey 3* suggest that few existing or prospective parents make much reference to KPIs, and it would seem unlikely that they will apply much direct pressure on schools. These findings would seem to supports those of Harrison (1995b) and Kouzim *et al* (1999) in chapter 2, and the Kirkland Rowell (2000) study[140] which place exam results as the 9th most important factor for parents when considering secondary schools in England (see also Woods *et al* 1998).

However, some care should be taken with reading too much in to this. There may be some parents who have, overtly or covertly, voted 'with their feet', and moved their children, and therefore their concerns over KPIs may not be apparent to the respondents. Furthermore, other cultural factors may be significant. In education generally, and perhaps more so in Scotland, schools and teachers are seen by many parents as respected and authoritative figures whose views and actions should not be questioned. And in a similar way some parents may simply not wish to make 'trouble' for their children.

---

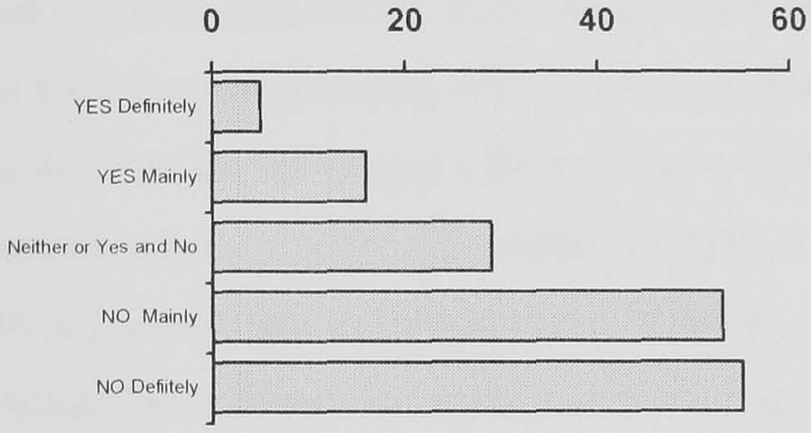
[140] Education perception monitor

Survey 1: Recall school's performance



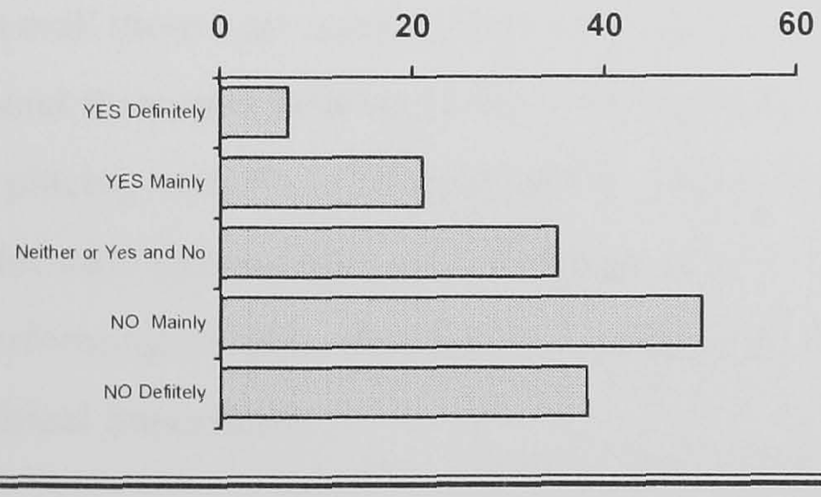
N=156 Mean=2.00

Survey 2: Existing parents reference to KPIs



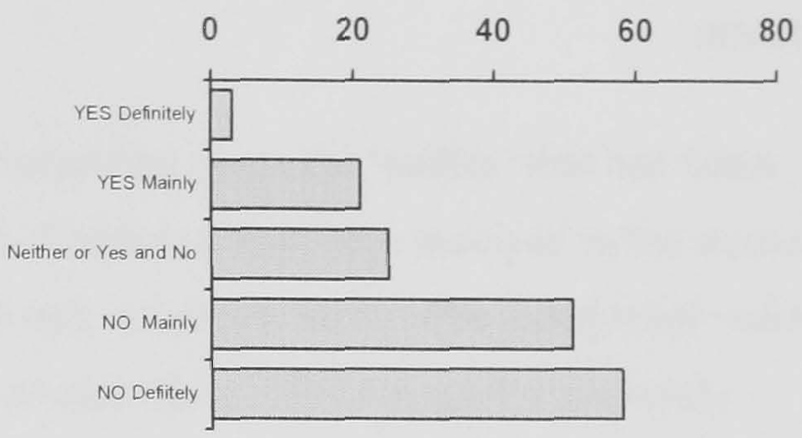
N=158 Mean=3.87

Survey 3: Prospective parents reference to KPIs



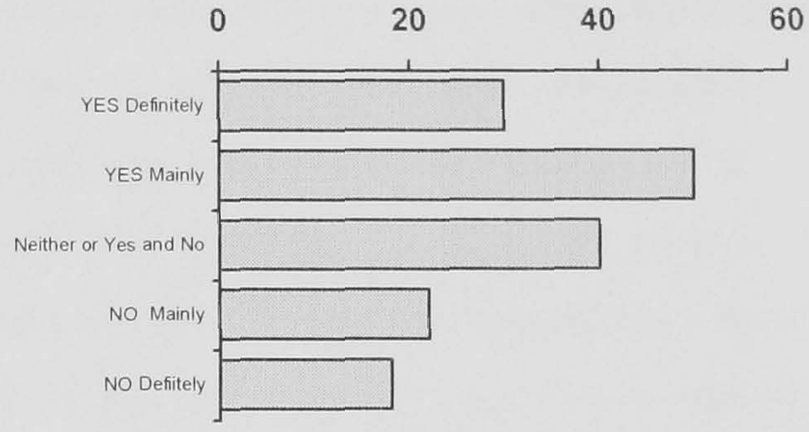
N=151 Mean=3.60

Survey 4: Aims in KPI terms



N=158 Mean=3.89

Survey 5: Importance of school's relative position



N=160 Mean=2.68

How do you feel about KPIs

These semantic differentials looked in general terms at the respondents' views of KPIs, *Survey 6* shows that the respondents generally found KPI systems slightly more beneficial than harmful; as a governor of a primary school pointed out;

> ....they (KPIs) provide a useful tool to help improve standards generally and in particular they are helpful in raising teacher/school performance. They are also helpful to governors in managing and planning school services...
>
> (RS 21)

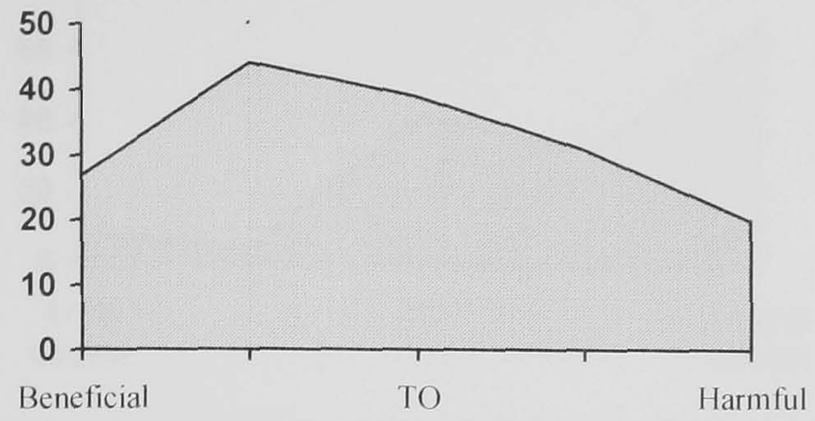
Nonetheless the level of support for KPIs was a little surprising given the 'battles' that had taken place since their inception (chapter 5) and the amount of criticism they have received in the media from schools, both north and south of the border; although in fairness most of the complaints were focused on individual aspects of KPI systems such as league tables. This response did broadly support the result from the preliminary questions (Chapter 10), namely that this sample of respondents are not inherently antipathetical towards KPIs and therefore it is argued should be able to provide a reasonably fair and balanced view.

Looking more specifically at the perceived effects of KPIs, *Survey 8* gives a very strong indication that they have a narrowing effect on the curriculum. There are echoes of the issues discussed in chapter 5 on the narrowing effect of the revised code, and later at the turn of the century of the grammar school system. This narrowing effect has been widely predicted, for example, Smith's (1995a) unintended consequences of publishing performance data. It might be expected that *Survey 7*, would in a similar way show that KPIs discourage innovation, however this was not so clear. It would seem likely that the question was at fault, with the problem being neatly summarised by RS (89); "do you mean innovation by finding new ways of passing the tests or doing new and innovative things in the curriculum".

Overall there was some evidence that the majority of respondents did not trust KPI systems and found them undermining (*Survey 10* and *Survey 9*). Given the reliance the government (chapter 5) is placing on KPIs to be used both to manage and hold accountable virtually all areas of the education systeim, these two findings are of concern. Indeed, whether for 'high' or 'low' performing schools, the issue of having trust and faith in the systems by which one is judged, is of critical importance.

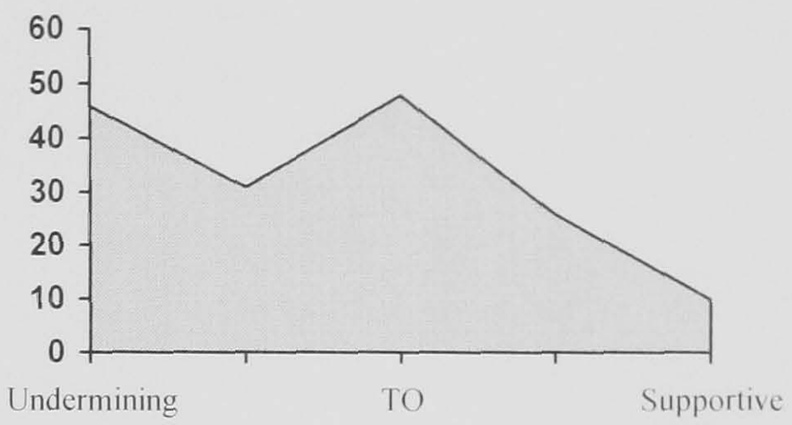
The issue of who controls schools was discussed in some details in chapter 2. In spite of governments claims that schools 'now' have more freedom (Morris 2001a) and local control, the respondents (*Survey 12*) strongly indicated that KPIs were used to increase central control. *Survey 11* aimed to give an indication of how prominent KPIs were in the respondents schools, and not surprisingly KPIs were for most, quite frequently discussed. This (and *Survey 1*) supports the notion that most respondents were very much aware of KPIs and in all probability aware of their effects.

Survey 6: Beneficial



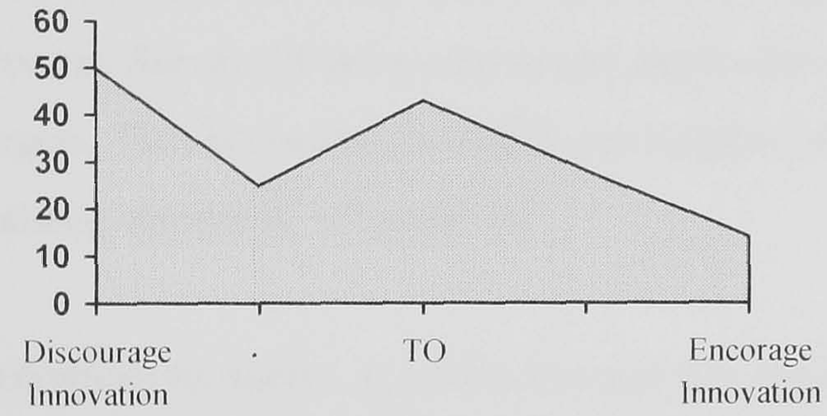
N=161 Mean=2.83

Survey 9: Undermining



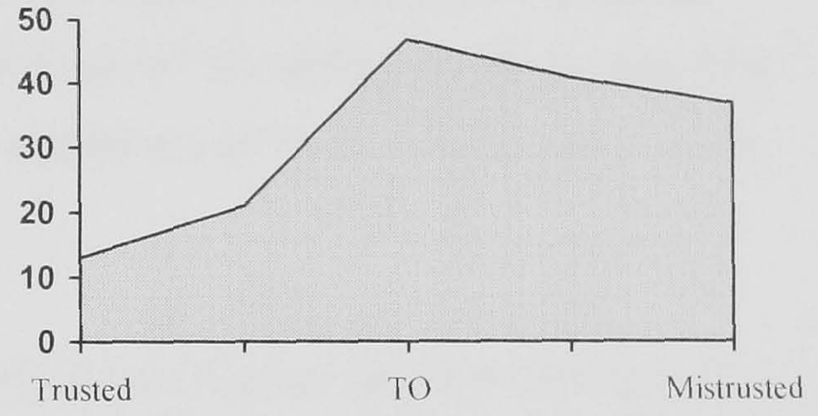
N=161 Mean=2.52

Survey 7: Innovation



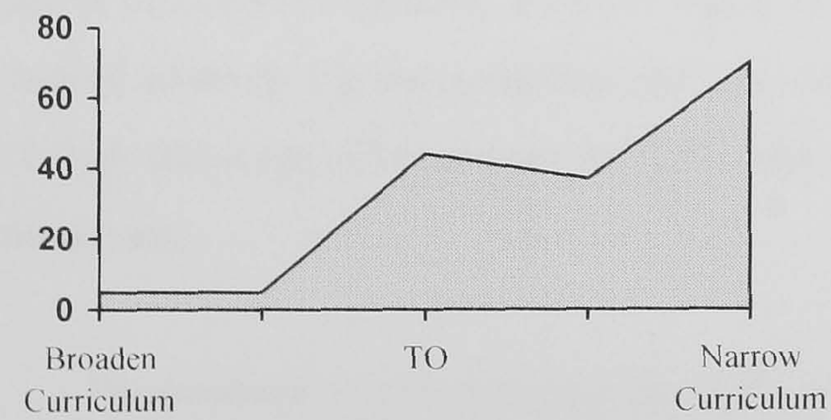
N=160 Mean=2.57

Survey 10: Trusted



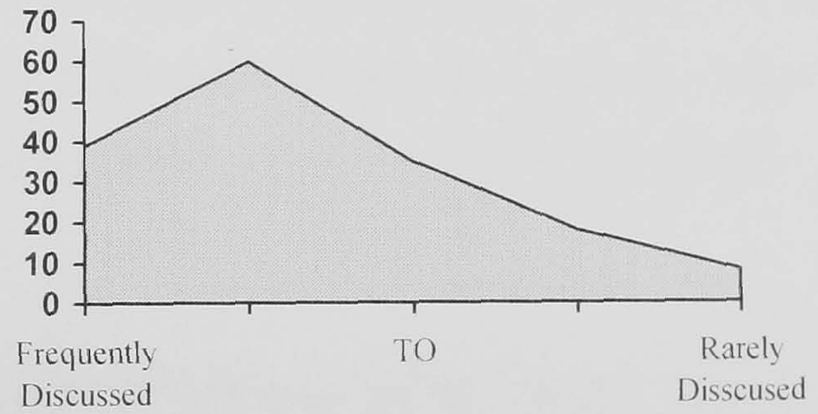
N=159 Mean=3.43

Survey 8: Curriculum



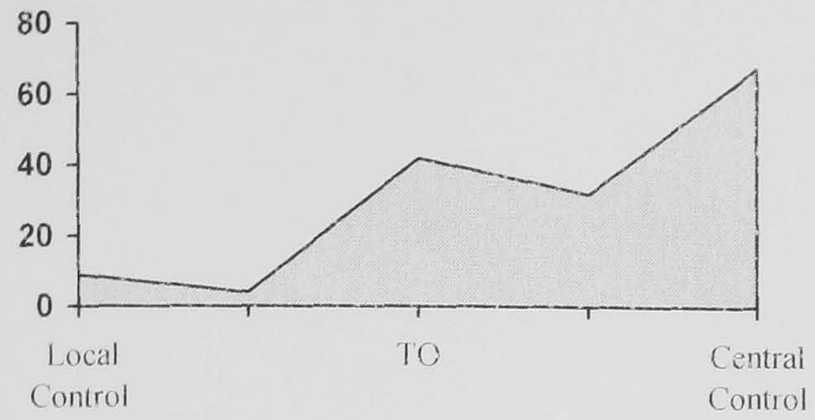
N=161 Mean=4.01

Survey 11: Discussed



N=160 Mean=2.35

Survey 12: Control



N=155 Mean=3.87

Targets and Target Setting

This section looks more specifically at the respondents' perceptions of targets and the setting process. *Survey 13* shows not surprisingly that the vast majority feel under pressure to meet their targets. This is consistent with the principles of NPM and the government's philosophy towards 'raising standards' (chapter 5).

In contrast to *Survey 4*, which showed that most schools did not express their aims in terms of KPIs, *Survey 14* found that most expressed their targets (or objectives) in KPI terms. Furthermore, the majority of respondents found little conflict between their aims and targets (*Survey 15*) and where there was conflict, *Survey 16, s*uggested that the targets did not tend to override the aims. However, conflict may not necessarily be bad, as Cowie (2000) points out, conflict can lead to creative and critical thinking. Even so much still depends on whether this thinking can be turned in to action. (*Survey 17*) indicated that the vast majority of respondents did not feel that they avoided activities that were not measured by KPIs. However, as a primary school serving a deprived area pointed out;

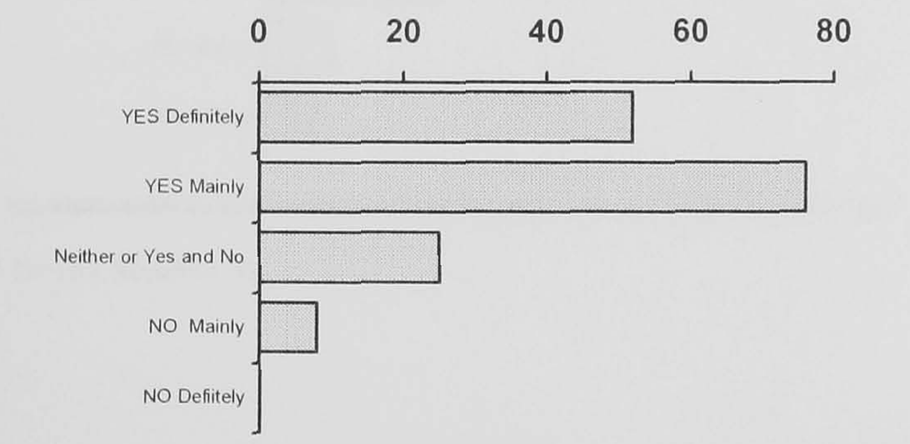
> The emphasis on only language and maths means that we are tending to focus too greatly on this. We are in an area of deprivation and skills relating to self-esteem and self-motivation are very important to us. These are not being reflected in National Test results.
>
> (RS 237)

Survey 18 showed strong support for the principle of setting targets on an individual pupil basis, rather than for the whole school. Making predictions at an individual level is far more meaningful than whole school 'guesstimates', which may well be based on crude formulas or have simply been handed down. Respondent RS 141 provided a good illustration of the pitfalls of simply factoring or ratcheting up last years target. His school (Grammar) had achieved 100% A*-C, however for the next year the LEA simply (and fairly?) added 5% to every school in the county, making his target a rather challenging 105%!

Not surprisingly *Survey 19* showed that the vast majority of respondents felt that they should participate in setting their own targets. Many of the key organisational theorists (eg. Mayo, McGregor) discussed in chapter 4, advocated a shared or participative approach to target setting. Indeed, the DfEE's school improvement model (*Figure 5*) essentially aims to adopt a joint approach. And *Survey 64* suggests that the majority of respondents do participate in the setting of their targets. However, as pointed out in chapter 5 there would appear to be moves to reduce school participation in favour of centrally imposed targets. Indeed as a Scottish secondary head (RS 254) pointed out; "Set by authority – no negotiation accepted". And an English counterpart;

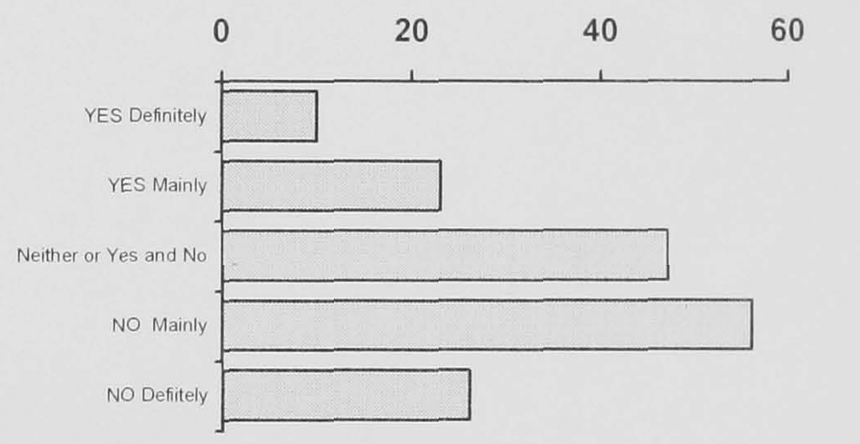
> The ludicrous mis-use of targets is a waste of my time: the government think up a figure that is 99% wishful political expediency. They give a 'target' to each LEA and lo and behold each LEA has decided it can reach its target in 4 Years in 4 even steps! This is pie-in-the-sky NONSENSE!! (RS 48)

Survey 13: Pressure to meet KPI targets



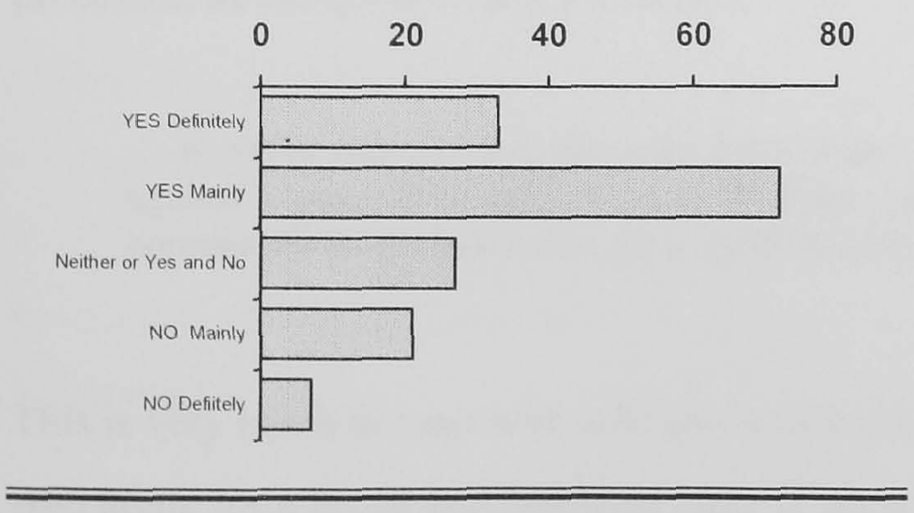
N=161 Mean=1.93

Survey 15: Conflict between aims and targets



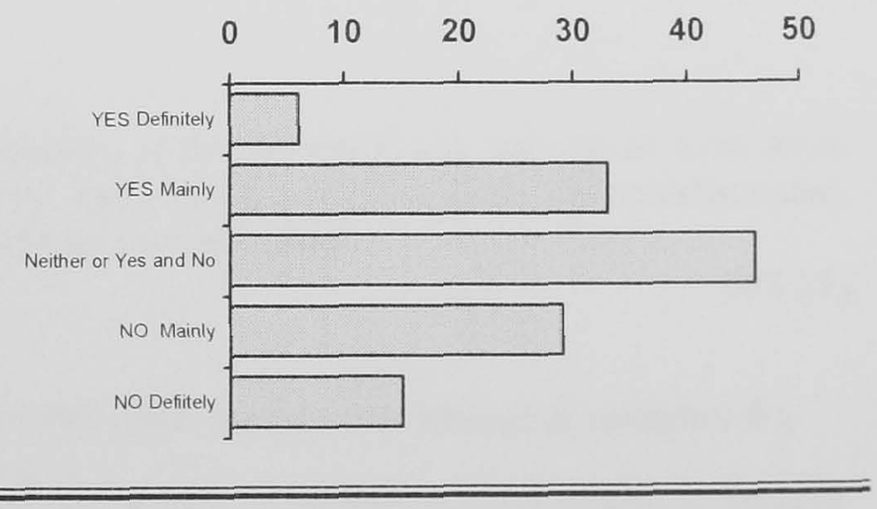
N=162 Mean=3.40

Survey 14: Targets expressed in KPI terms



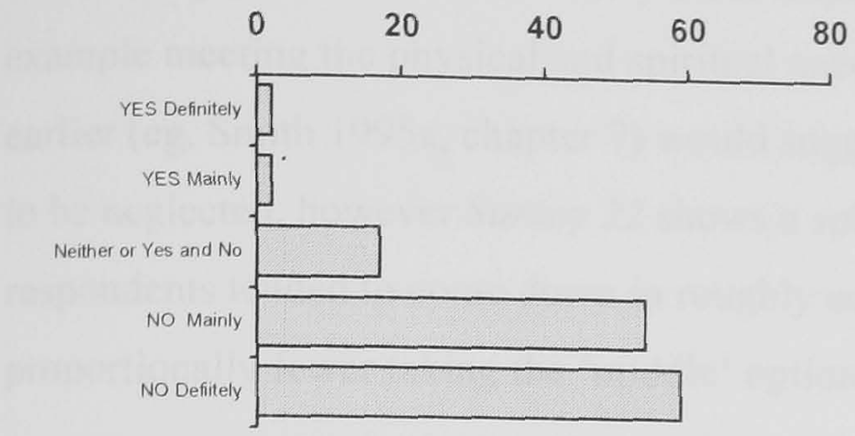
N=160 Mean=2.36

Survey 16: Targets override aims



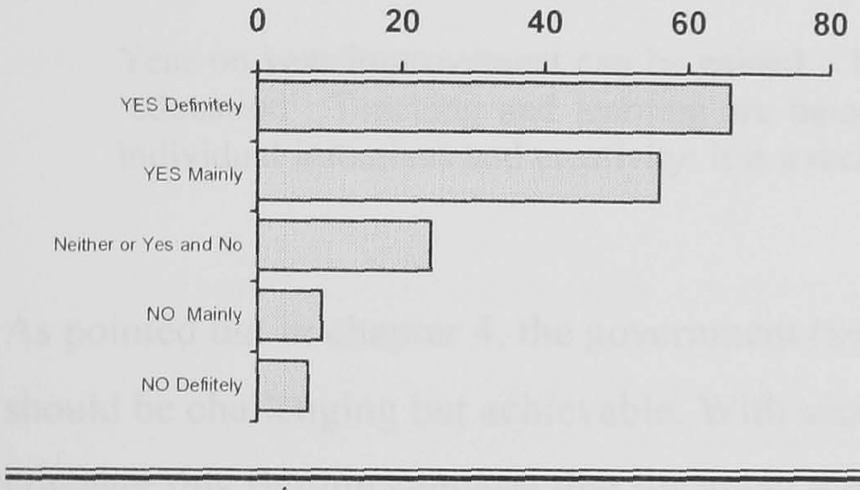
N=129 Mean=3.11

Survey 17: Avoid activities not measured by KPIs



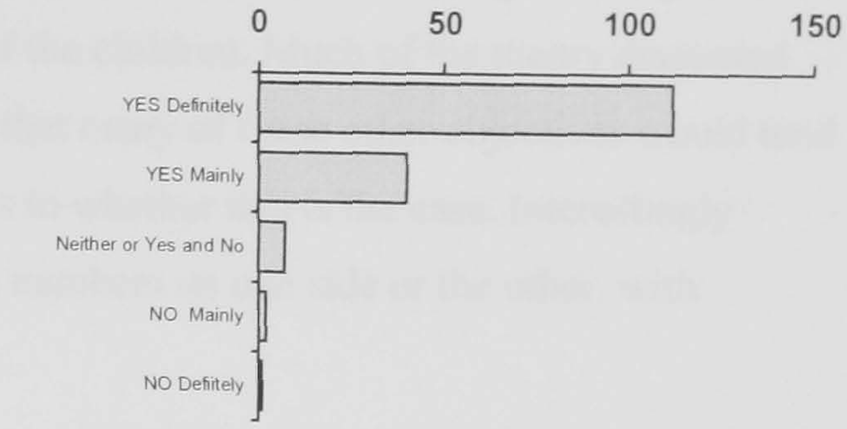
N=134 Mean=4.29

Survey 18: Individual pupil targets



N=162 Mean=1.98

Survey 19: Set own targets



N=162 Mean=1.40

Statutory targets – more specifically

This section looks in more detail at statutory or KPI targets. The first question, *Survey 20,* shows overwhelming support for the notion that KPI targets should consider more aspects of school processes, as the quote below illustrates;

> ....however they are not the sole determinant / indicator of performance and need to be considered against a number of other factors including pastoral care, special need support, team commitment, community involvement and the overall ethos/performance of the school.
>
> (RS 21)

This is very much in tune with Johnson and Kaplan (1991) and many other theorists (chapter 6), who argue for a broad and balanced view of performance measures; indeed a view which is largely accepted by some parts of the government (chapter 2). In practice this is closely related to the principle of a 'balanced score card' which is discussed later in the conclusion.

There was a split over whether the statutory targets help raise overall performance (*Survey 21*). Much of the underlying philosophy of using a limited number of KPIs is that in addition to

improving the performance of what they measure, other aspects will also improve: in effect raise the whole game. Schools have many other important objectives besides meeting KPI targets, for example meeting the physical and spiritual needs of the children. Much of the theory discussed earlier (eg. Smith 1995a, chapter 7) would suggest that many of these other objectives would tend to be neglected, however *Survey 22* shows a split as to whether this is the case. Interestingly respondents tended to come down in roughly equal numbers on one side or the other, with proportionally fewer taking the 'middle' option.

Survey 23 showed that although most schools were confident that their KPI targets would be achieved, a significant number were not so sure. And as an English secondary points there may well be 'costs' attached;

> Year on year improvement can be gained – but at what cost to what many consider the wider aims of 'education'. Teaching and learning are being placed in a straightjacket with less and less room for individual initiatives and creativity: it is a recipe for social paralysis.
>
> (RS 20)

As pointed out in chapter 4, the government (supported by conventional wisdom) states that targets should be challenging but achievable. With such an approach it is inevitable that a few will miss, however this finding suggests that for many this may well be a serious issue. Much of the problem may well be due to globally imposed targets which do not take account of the natural variations in the cohorts (chapter 2).

Leading on from this *Survey 24* shows a split between those that feel that the target setting process makes appropriate allowances for their circumstances, and those that do not. A number of specific examples were given by the respondents as to how the system was not fair by making allowances. At the school level, RS (47) a secondary head of a grammar school in Kent (high in the league tables) was somewhat exasperated by his governors; "I am tired of governors using KPIs without taking local factors on board". Maybe this is an isolated example, or not what it seems, but it must be of concern if a head (in this relatively comfortable position) does not have faith in his governors' abilities to interpret KPI data. Interestingly he was more sanguine about parents judgements.

The most frequently cited reason for unfairness was that the KPI systems did not make allowances for special needs pupils. Whilst this may well be unfair for the particular school, perhaps the most worrying aspect (not mentioned by any of the respondents) would be schools avoiding such pupils. Another issue highlighted by a number of respondents was the problems of small schools, for example (RS 4); "Being a very small school statistics quoted as percentages are totally

meaningless. 17% can indicate one child's performance". The fact that the 17% can be either way accentuates the problem[141]. An all boys school (RS 49) made an interesting point that because of girls outperforming boys in terms of 5 A*-C it was unfair to make comparisons with mixed and all girl schools. Another head (RS 98) suggested that KPIs should take account of differentials in funding. As discussed in chapter 2 such proposals exist (Mayo 2000) but whether they would be any fairer is very questionable.

Together these issues raise questions about the target setting process, in particular the role of the (L)EAs, which has appreciably changed over the last few years (chapter 5). However, rather surprisingly *Survey 26* does not particularly support this assertion that the relationship has changed, with the majority of respondents being neutral. Research by Cowie (2000) on the relationship between secondary heads in Aberdeen and their EA highlighted the different perceptions, and significantly pointed to the consensus among the heads that a new relationship should be developed;

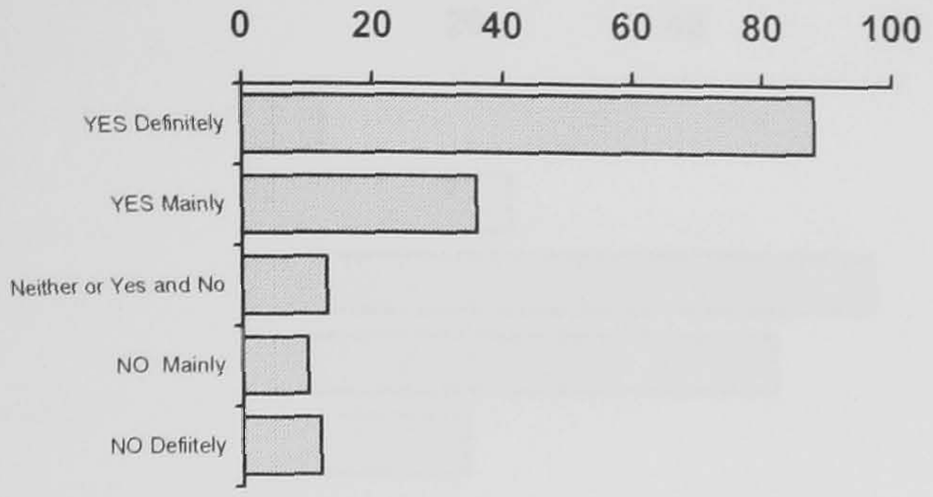
> The interview data also suggests a mismatch between the role of the headteacher as perceived by the headteachers and the Directorate, and considerable consensus emerged around the kind of relationship headteachers believe the Authority should develop with schools, supporting a view that there may be potential in developing a new role for headteachers in the governance of secondary schooling.
>
> (p.17)

However, he points out that what might be desirable for heads and perhaps EAs, may not be acceptable to central government. In spite of the key role of (L)EAs in the process of target setting, a large majority of schools indicated that they did have access to suitable information to set their own targets (*Survey 25*). Several respondents commented on use of third party systems, such as those from their LEA or Durham CEM, which they found useful for setting meaningful pupil level targets.

---

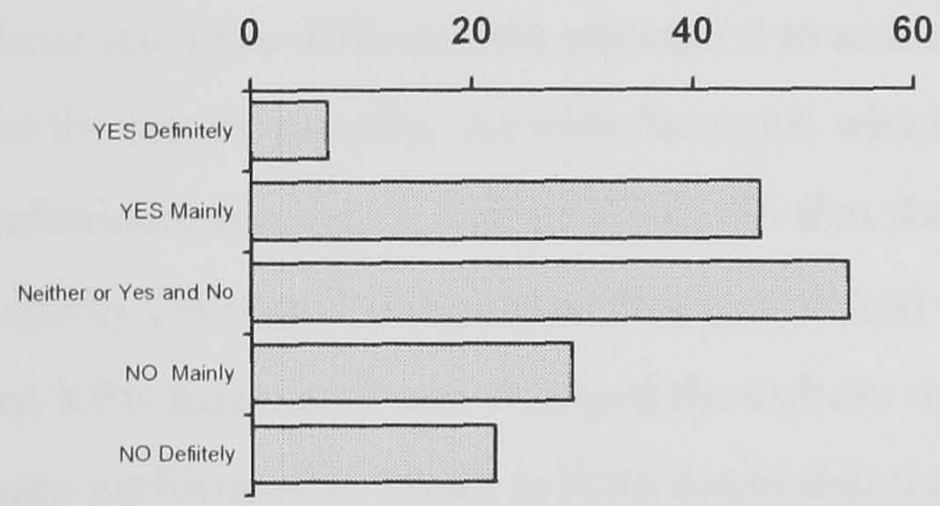
[141] It should be noted that in this case the results would not be made public, but of course they would still be used by others such as the governors and LEA.

Survey 20: More aspects



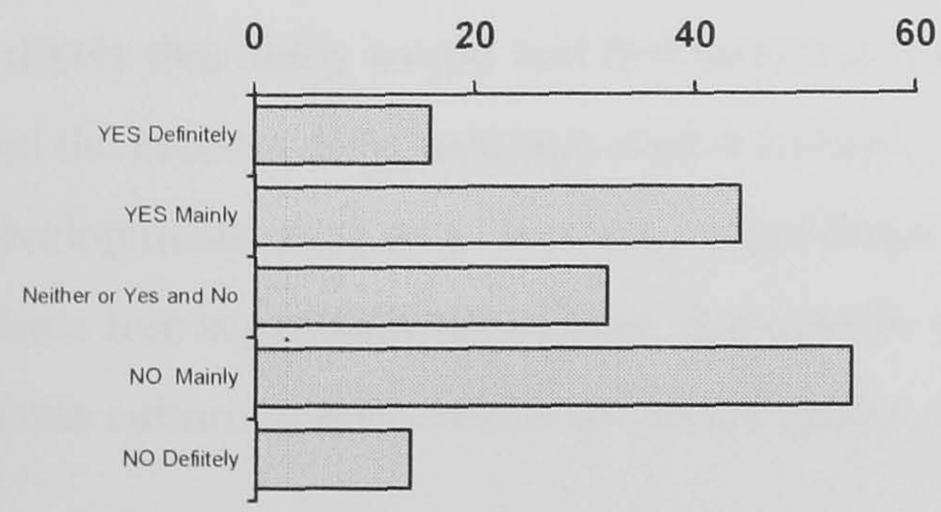
N=159 Mean=1.88

Survey 21: Raise overall standards



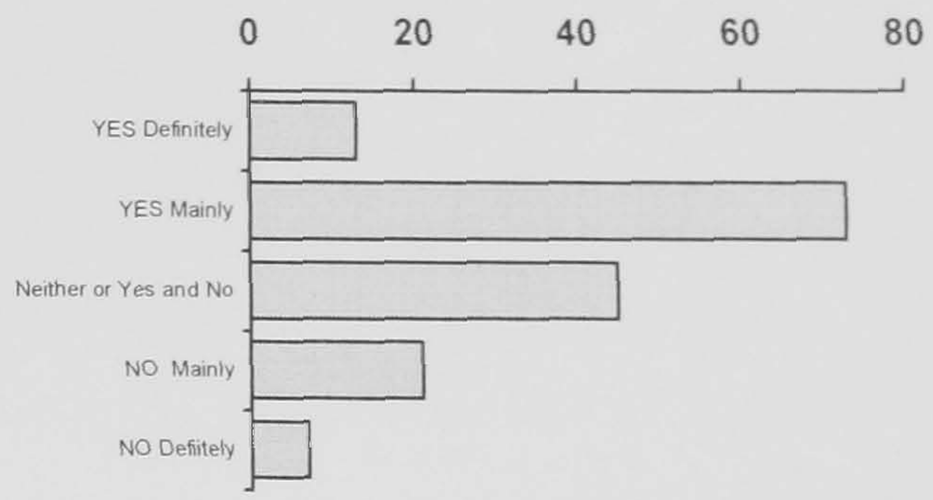
N=158 Mean=3.08

Survey 22: Other important objectives



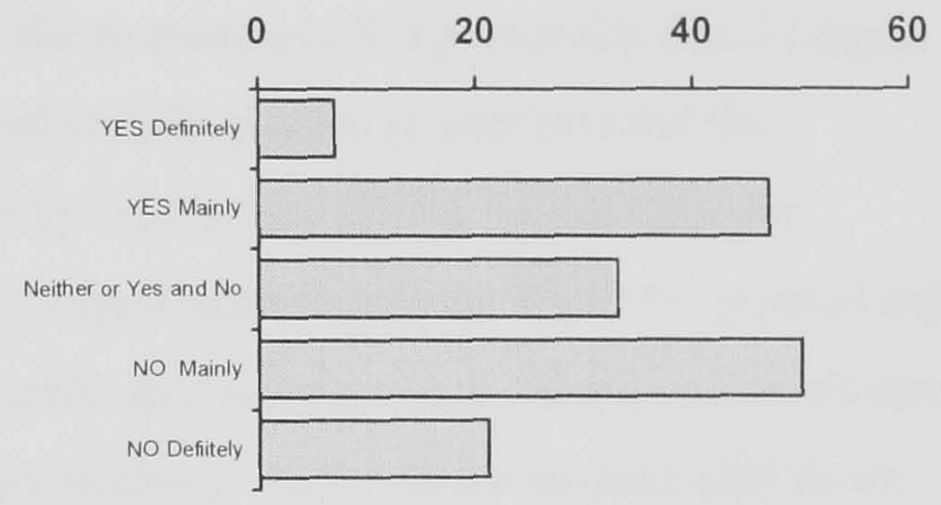
N=160 Mean=3.04

Survey 23: Confident targets will be achieved



N=159 Mean=2.60

Survey 24: Targets allow for particular circumstances

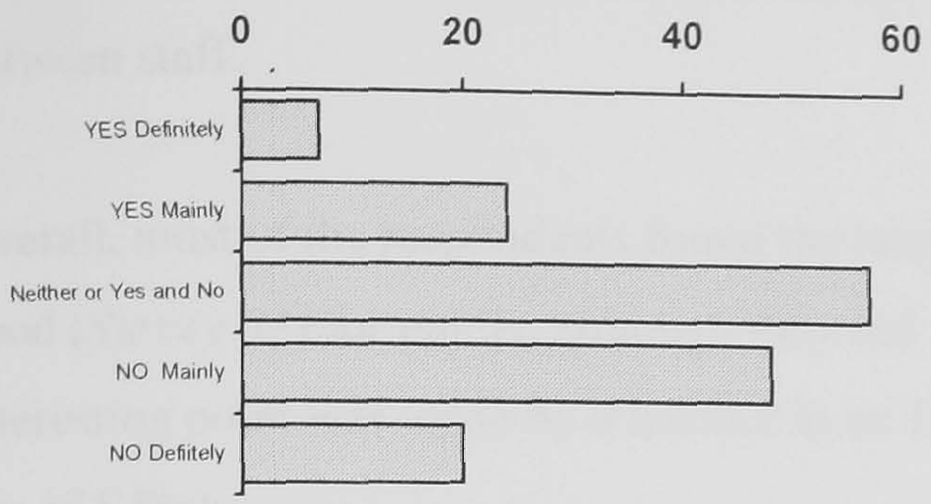


N=158 Mean=3.20

Survey 25: Suitable information for target setting

N=160 Mean=2.46

Survey 26: Changed relationship with (L)EA



N=156 Mean=3.32

Your feelings on Targets and target setting

These semantic differentials attempted to assess how the respondents felt personally about targets and the setting process. As with *Survey 6* which looked at KPI systems in general (and the preliminary research), *Survey 27* shows that the majority find targets useful. In addition the majority (*Survey 28*) indicated that they found targets helped raise standards. RS (141) pointed out that KPIs had raised and changed the culture in a positive way, to one which valued and celebrated exam performance. There is little doubt that targets are useful, however there are potential down sides.

A number of theorists, (eg. Argyris chapter 6) have argued that target setting (budgeting) can increase the blame culture in an organisation; *Survey 29* very much supports this. Whilst it was unlikely that many would feel that targets actually reduce blame, a very large proportion indicated that the process did encourage such a culture. And this would seem likely to hinder (desirable) developments, such as a 'learning organisation' approach (chapter 4) which relies on a relatively blame free organisational culture. It would be wrong however to necessarily attribute all of the blame culture to KPIs which are made public. An English secondary head observed;

> I have seen teachers in individual 'low performing' subjects as measured by ALIS criteria-become very defensive – yet there may be valid reasons for the exam results for a particular group.
>
> (RS 81)

She went on to point out the need to use tact and care with such indicator systems. This very much supports the notion that it is necessary to look to the broader organisational and cultural issues when assessing the impact of any indicator system.

There was also some evidence (*Survey 33*) that the target setting process worsens the relation ship between staff, although as RS (34) pointed out, much depends on whether the results are 'good' or 'bad' for that year. However, RS (77) took a more extreme view; "We are living in a fear regime which is counterproductive to education in its true sense". This may have a significant effect on

other aspects of the system. For example, many of the key challenges facing education, such as the government's inclusion agenda, require a high degree of cooperation, trust and good relationships between staff.

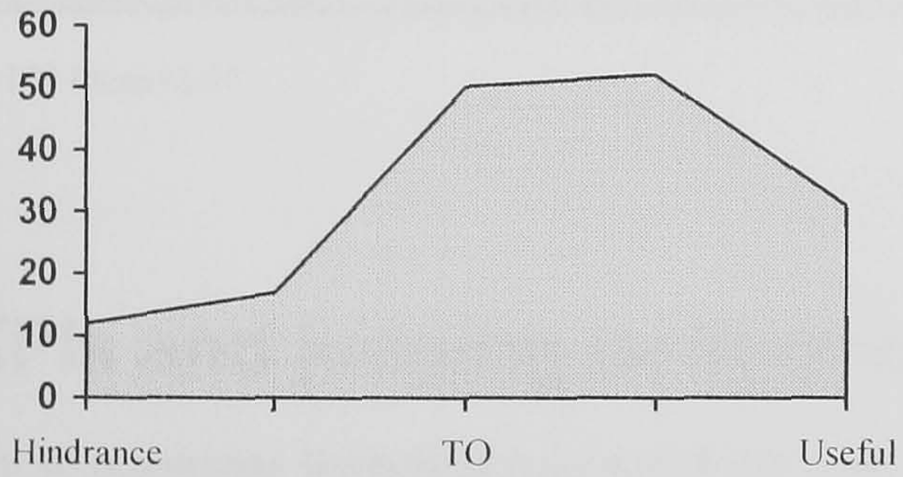
Overall, most of the respondents found the target setting process was both relevant (*Survey 30*) and good (*Survey 31*) for pupils, although they did find them a little restrictive (*Survey 32*). An interesting point was made by a teacher in an English secondary school (who broadly supports the use of KPIs);

> ....however students feel manipulated by the increasing emphasis on results: they feel 'used' and that they can be regarded as cogs in a machine.
>
> (RS 69)

In a similar way another secondary school found that;

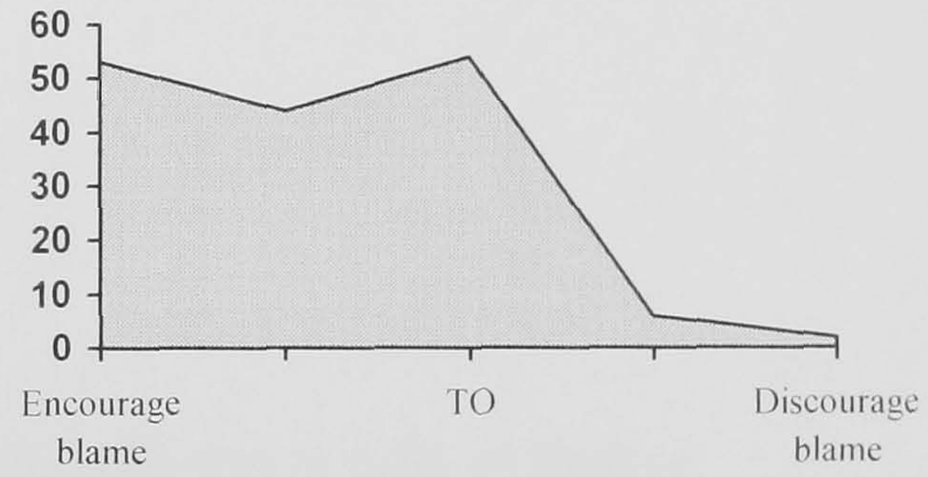
> Above all they increase the trend for students to study more and more subjects to increase point scores.
>
> (RS 47)

Survey 27: Hindrance



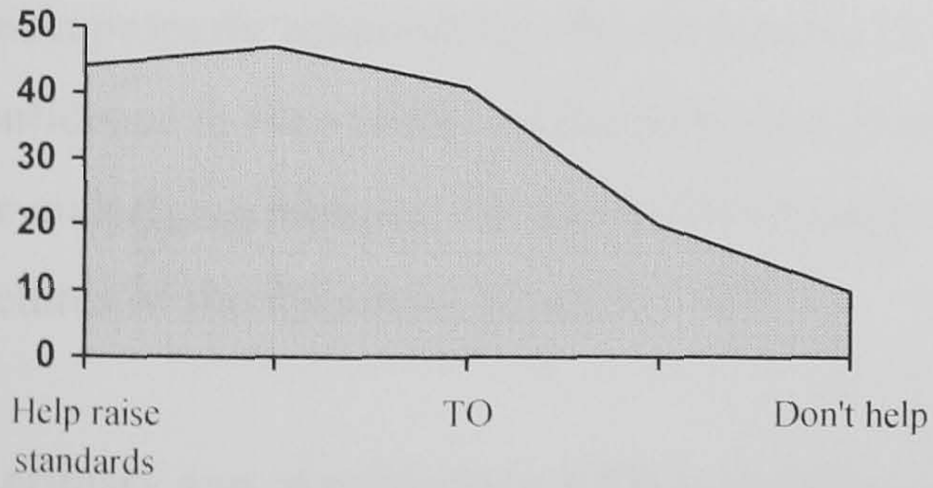
N=162 Mean=3.45

Survey 29: Blame Culture



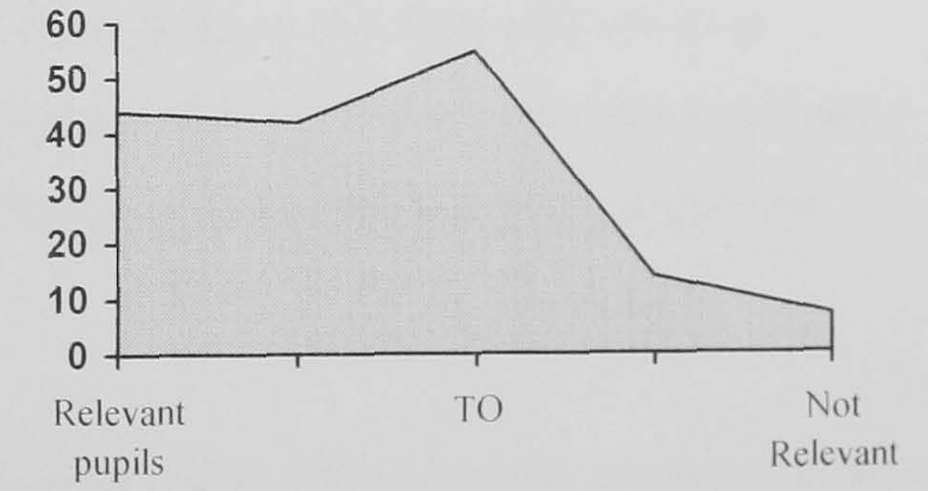
N=159 Mean=2.12

Survey 28: Raise standards



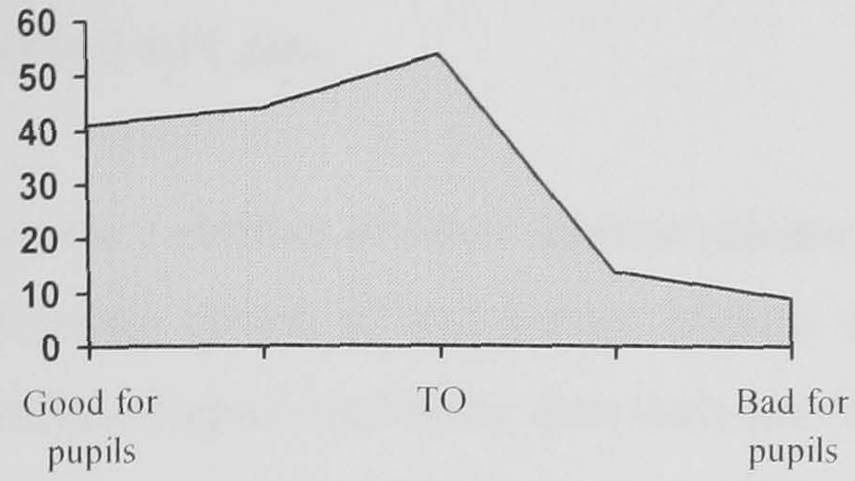
N=162 Mean=2.41

Survey 30: Relevant



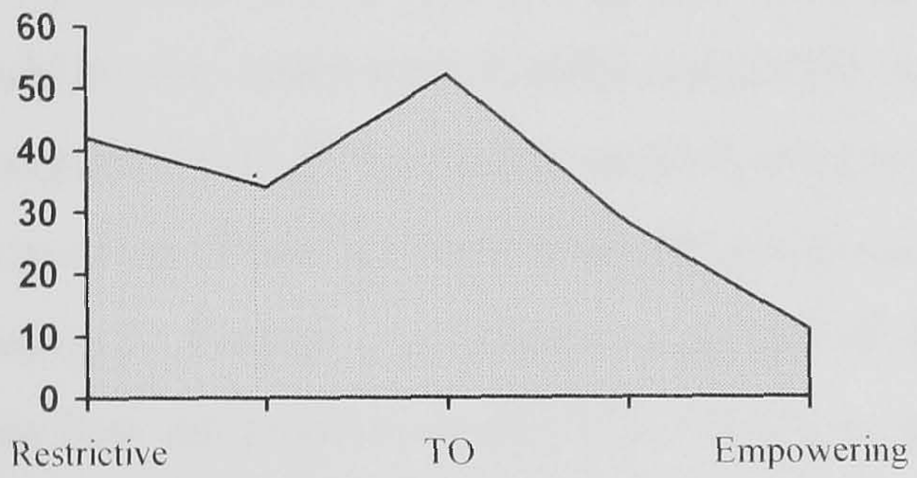
N=162 Mean=2.37

Survey 31: Good for pupils



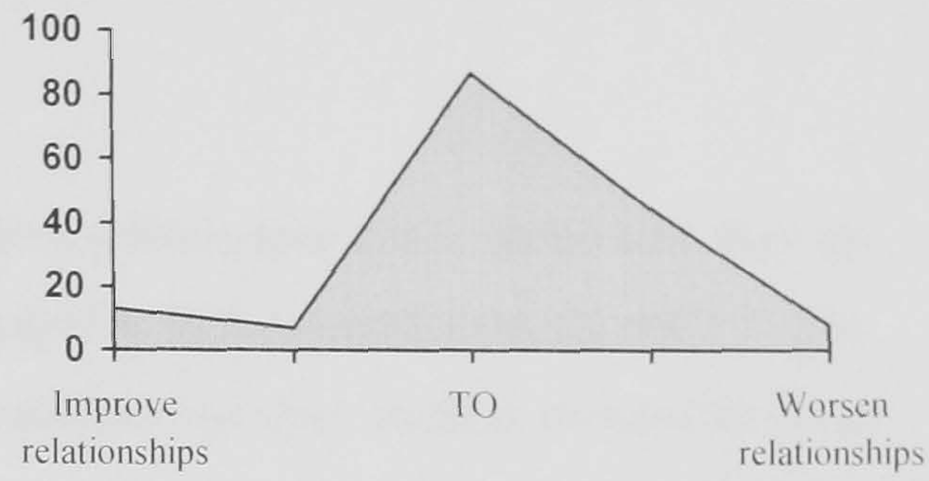
N=162 Mean=2.42

Survey 32: Restrictive



N=159 Mean=2.67

Survey 33: Relationship staff



N=160 Mean=3.18

KPIs and judging performance

These questions looked at how valid the respondents feel KPIs are in terms of judging performance. The concept of holding schools accountable by publishing their KPIs has been central to the government's philosophy since the citizens' charter (chapter 5). And although official tables may be obtained by the public, they are mainly promulgated through the media. *Survey 34* illustrates a rather worrying perception that the vast majority of respondents do not feel that the media properly acknowledge the limitations of KPIs, and it follows that they will not have confidence in the validity of the published results. For some (RS 7) there are grounds to question the underlying motives; "Selective information is attractive to the media but gives inaccurate pictures of the dynamics of school life".

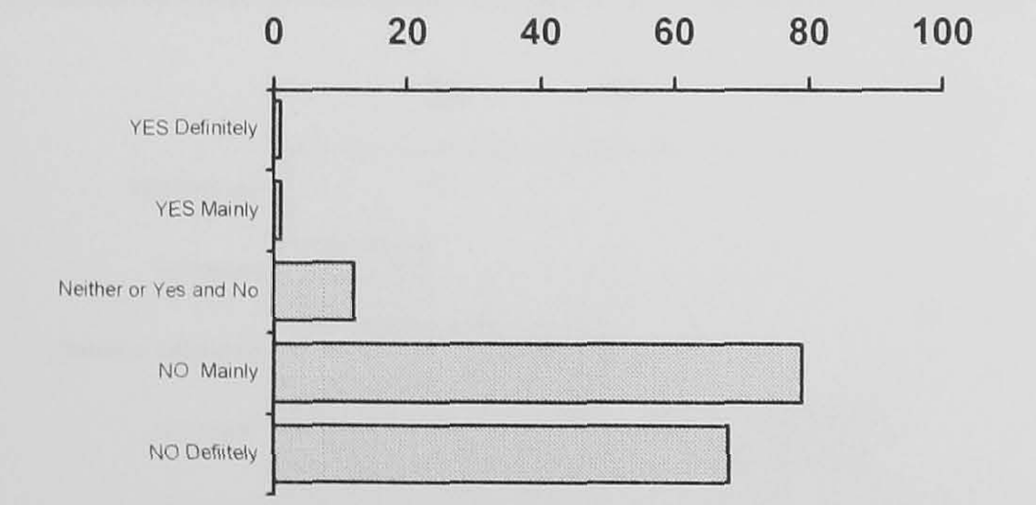
A primary aim of publishing KPIs is to help parents decide which school is most suitable for their children; however *Survey 35* shows that the overwhelming majority of respondents feel that they do not do this. Leading on from this *Survey 36* shows that most respondents do not feel that parents understand the limitations of KPIs. Indeed, a primary school (RS 70) noted that; "We have begun

to have meetings to 'explain' league tables to parents". Taken together these findings, question the government's view (chapter 1, Tomlinson 2001) that parents are able to properly interpret published KPI data.

There are a number of other sources besides KPIs which schools may use to judge how they are performing. *Survey 37* and *Survey 38* show that the majority of the respondents do not feel that KPIs have higher credibility than their own internal indicator systems, or their own professional judgements. As RS (250) pointed out; "National tests are used to confirm the teachers judgement". Whilst this would seem a positive finding, the obvious difficulty is if these different messages conflict.

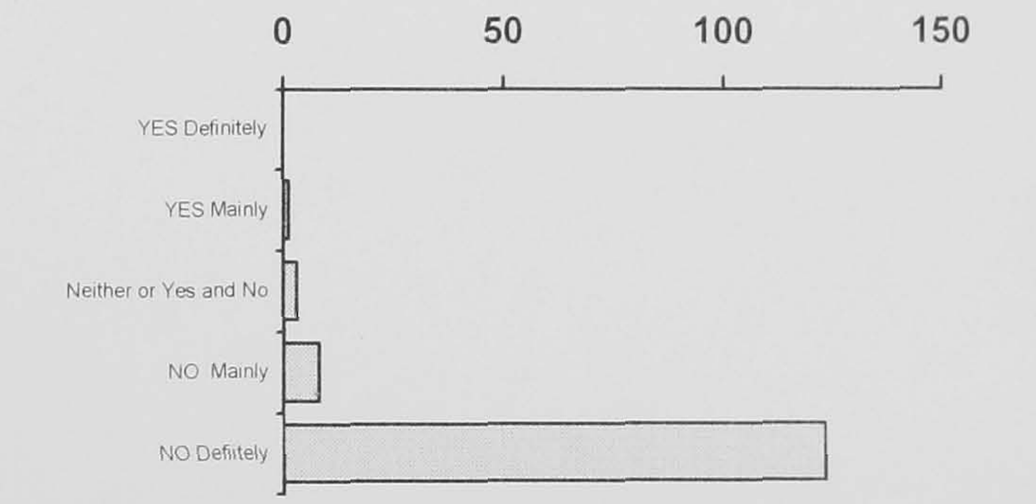
Survey 39 and *Survey 40* only apply to Scottish primary schools. These show that the respondents would be very much against publishing KPIs, and very little support for formalising the testing arrangements (ie. from teacher administered to exam based). These findings largely confirm that the position of teachers and heads has remained consistent since such testing was first mooted (chapter 5). Perhaps a reasonable criticism of the initial objection might be that the teachers did not know how the process would work - now they have been able to see their English counterparts experiences; their objections remain.

Survey 34: Media acknowledge limitations



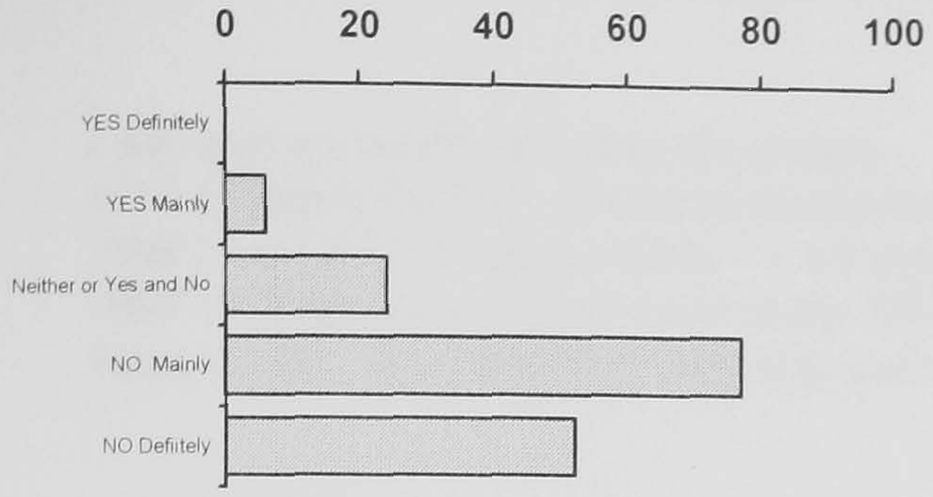
N=161 Mean=4.32

Survey 35: Indicate most suitable school



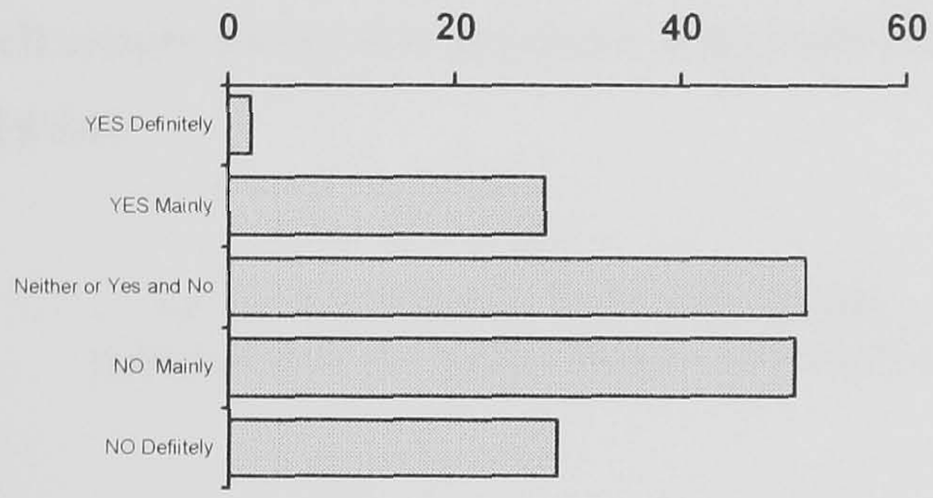
N=135 Mean=4.87

Survey 36: Parents understand limitations



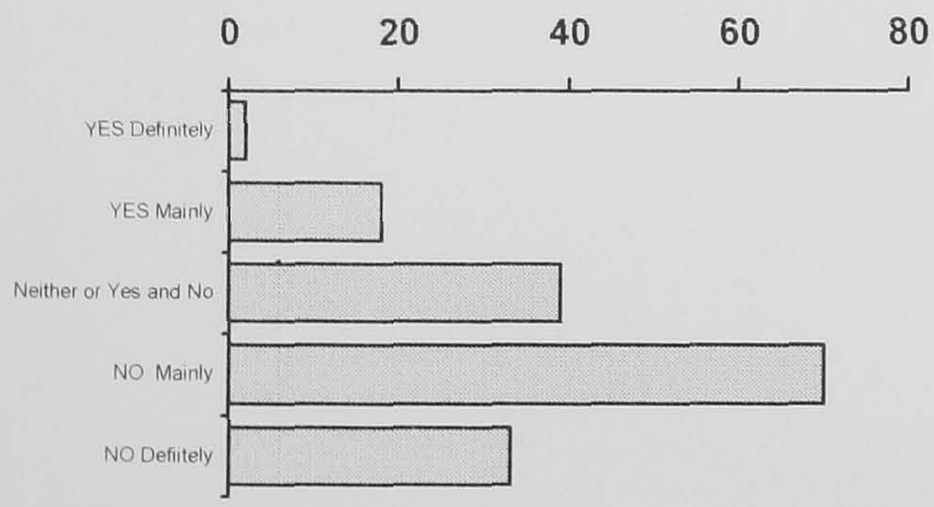
N=159 Mean=4.10

Survey 37: Higher credibility than internal indicators



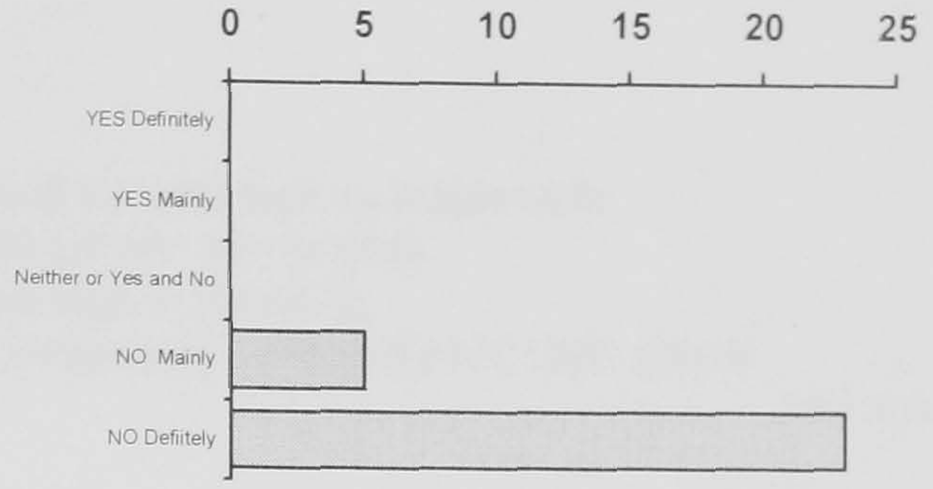
N=160 Mean=3.48

Survey 38: Higher credibility professional judgements



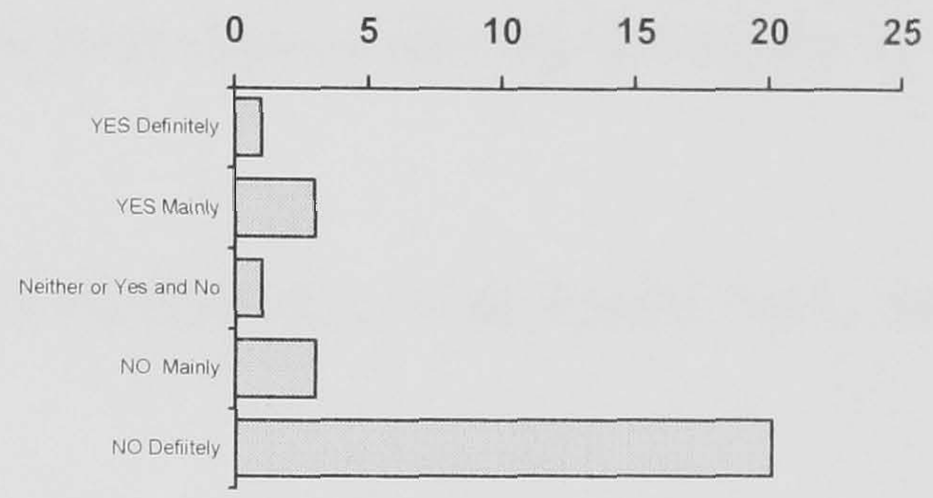
N=162 Mean=3.70

Survey 39: National test results made public



N=28 Mean=4.82 Scottish Primary only

Survey 40: More formal National testing



N=28 Mean=4.36 Scottish Primary only

Individual KPIs judging performance - England

As pointed out in the introduction due to the technical differences between Scottish and English KPIs the results from these questions are reported separately. Questions for *Survey 41* to *Survey 45* asked how valid individual KPIs are at judging school performance. These were asked of all English schools (primary and secondary) unfortunately quite a few of the respondents only commented on their own phase. The main indicator for primary schools, the number of KS2 level 4

and above was found to be moderately useful. The instability problem of this KPI was discussed in chapter 2, and was highlighted by a respondent;

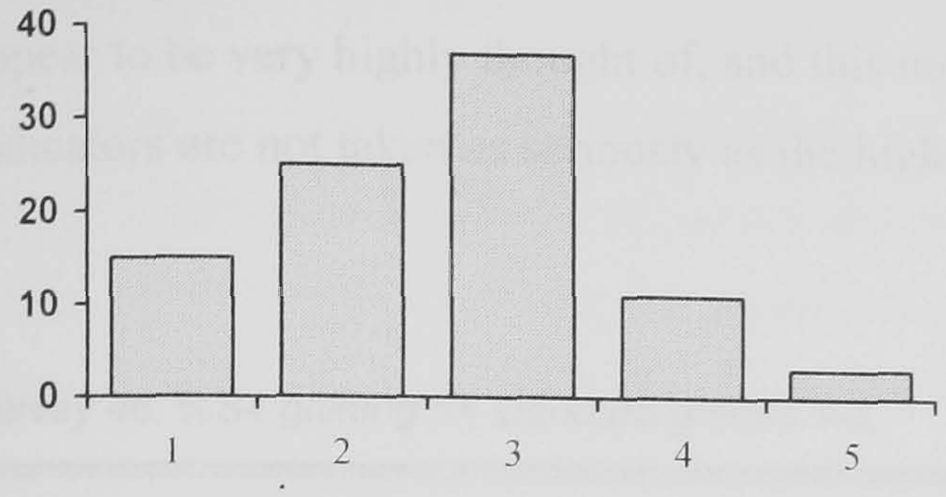
> I will indicate briefly why they are useless
> 1997 13 pupils for KS2, (1 special needs) result 91% at L4 very high up league table
> 1998 13 for KS2 (5 special needs + 1 L4 absent) 50% L4 very low in table
> 1999 19 (2 special needs) forecast result 78%[142] again high in the tables
> Same teacher, same teaching, same approach, same preparation – DIFFERENT CHILDREN
>
> (RS 104)

At the secondary level, the number of pupils gaining just one bottom level (G) pass was found to be least useful, followed by the percentage of pupils gaining 5 A* to C's. The point score was found to be better, and the value added, the most useful. On the face of it this would seem to be a good argument for publishing this in place of the 5 A* to C's. However, as argued in chapter 2 this might well simply create new problems and potential dysfunctional effects; a point largely supported by (RS 64):

> ....we as a governing body rely greatly on value added statistics. We do however believe that publication to the public would not benefit us

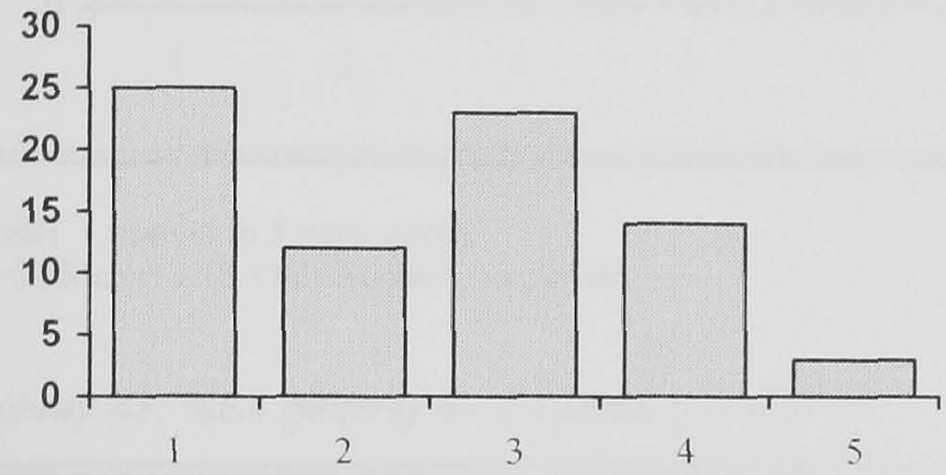
[142] The actual results were slightly better at 89% - ie. 2 pupils

Survey 41: KS2 level 4



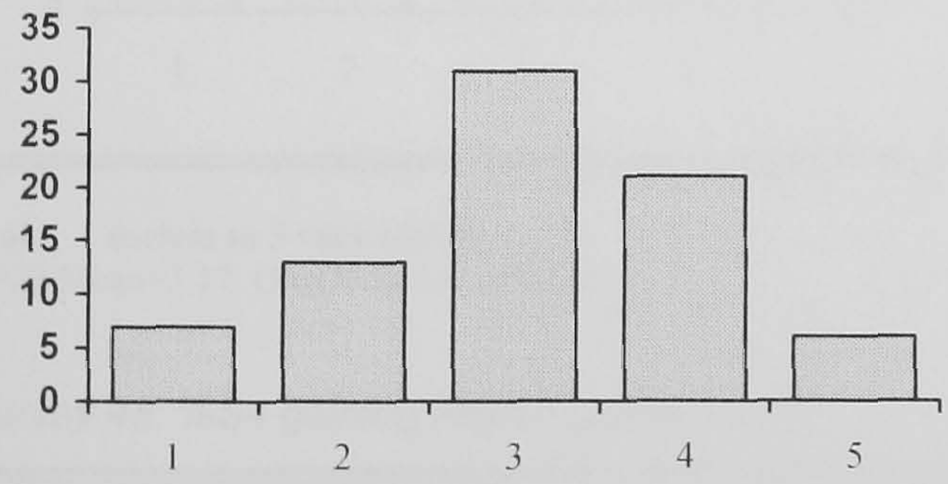
Scale: 1 useless to 5 very useful
N=91 Mean=2.76 Only English Schools

Survey 42: GCSE one G pass



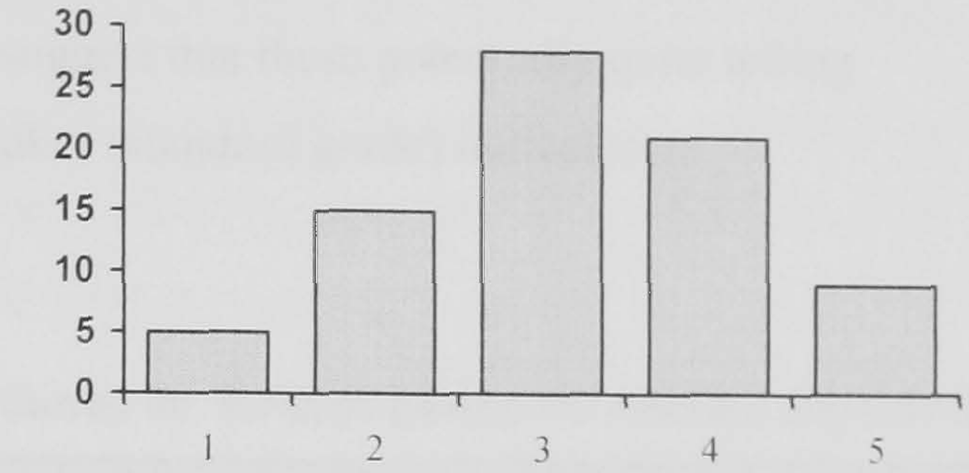
Scale: 1 useless to 5 very useful
N=77 Mean=2.62 Only English Schools

Survey 43: GCSE 5+ passes



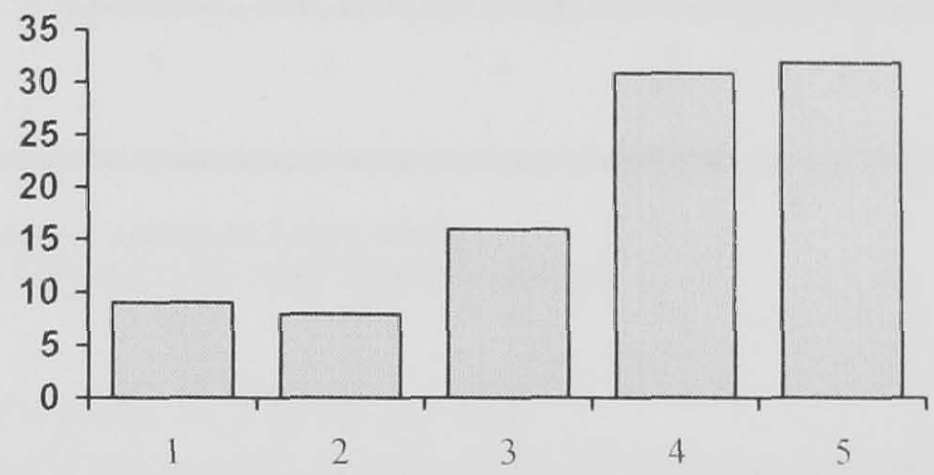
Scale: 1 useless to 5 very useful
N=78 Mean=3.08 Only English Schools

Survey 44: GCSE point score



Scale: 1 useless to 5 very useful
N=78 Mean=3.11 Only English Schools

Survey 45: Progress or value added measure



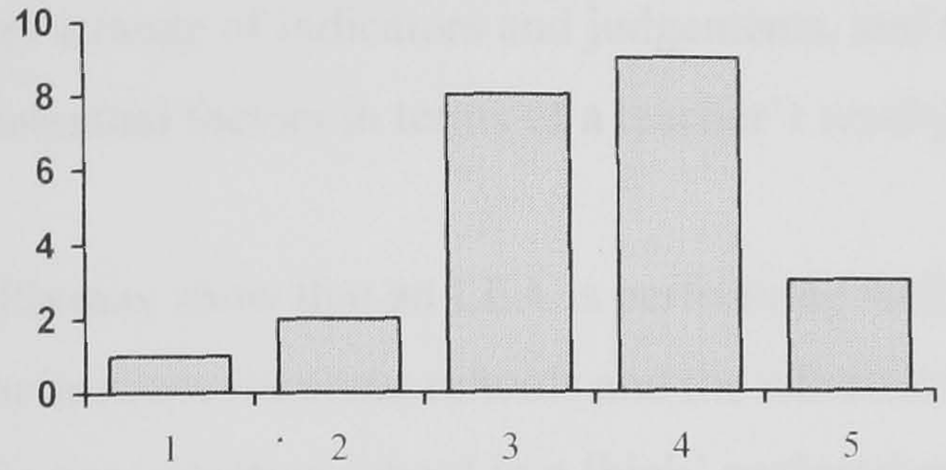
Scale: 1 useless to 5 very useful
N=96 Mean=3.52 Only English Schools

Individual KPIs judging performance – Scottish Secondary

Survey 46 to *Survey 50* give the responses from Scottish secondary schools towards their KPIs. Given the relatively low number of cases care should be taken with these results. Like England it would appear that higher levels of performance (ie. Grades 1-2) are considered more useful judges

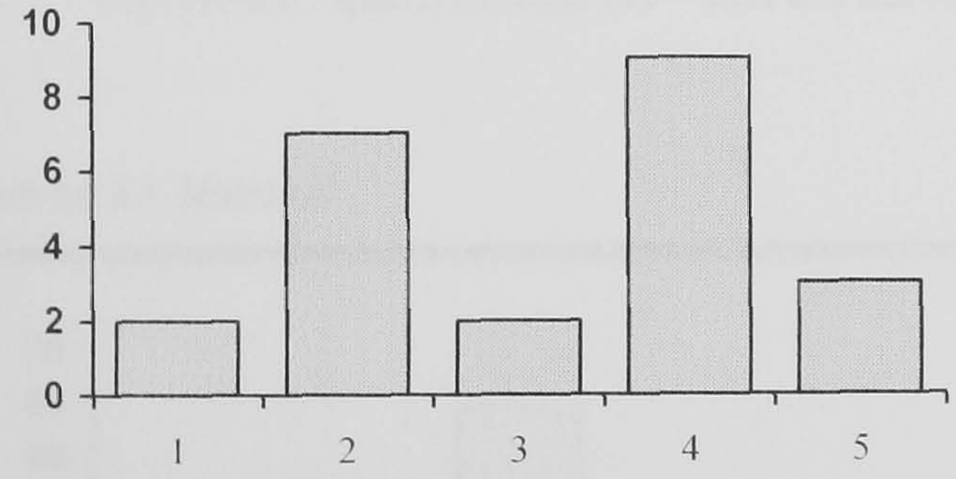
of performance than 'lower' performance (ie. Grades 1-6 and vocational qualifications). Slightly surprisingly the two indicators of possible future performance (*Survey 48* and *Survey 50*), do not appear to be very highly thought of, and this might suggest that these potentially quite telling indicators are not taken as seriously as the higher stakes (standard grade) indicators.

Survey 46: %S4 gaining 5+ standard grades 1-2



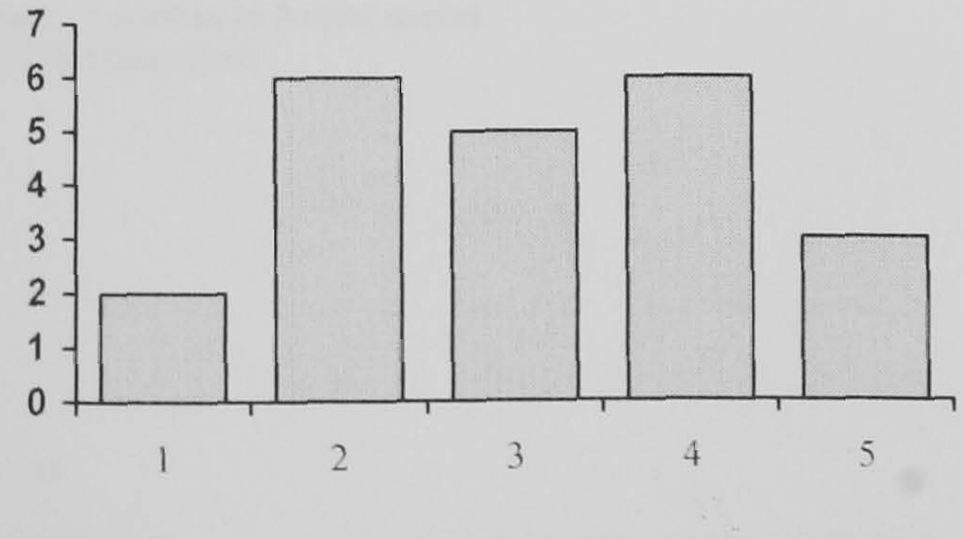
Scale: 1 useless to 5 very useful
N=23 Mean=3.48 Only Scottish secondary

Survey 47: %S4 gaining 5+ standard grades 1-6



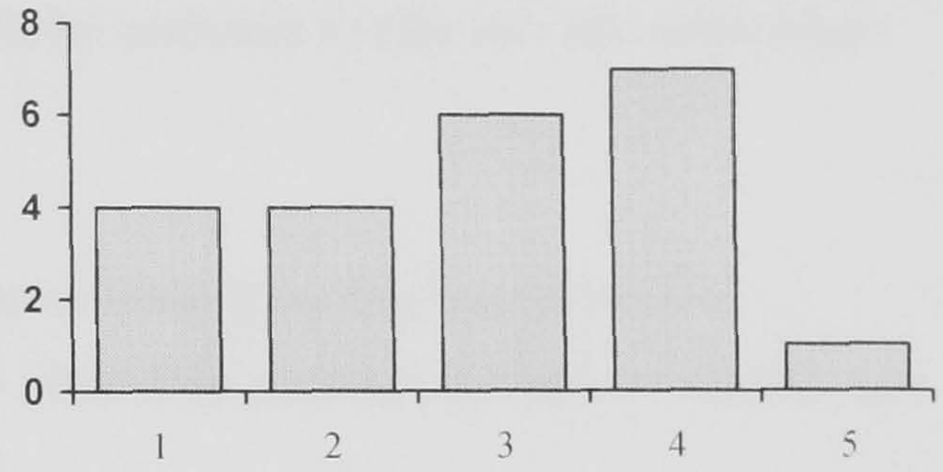
Scale: 1 useless to 5 very useful
N=23 Mean=3.17 Only Scottish secondary

Survey 48: %S4 gaining Higher grades in S5



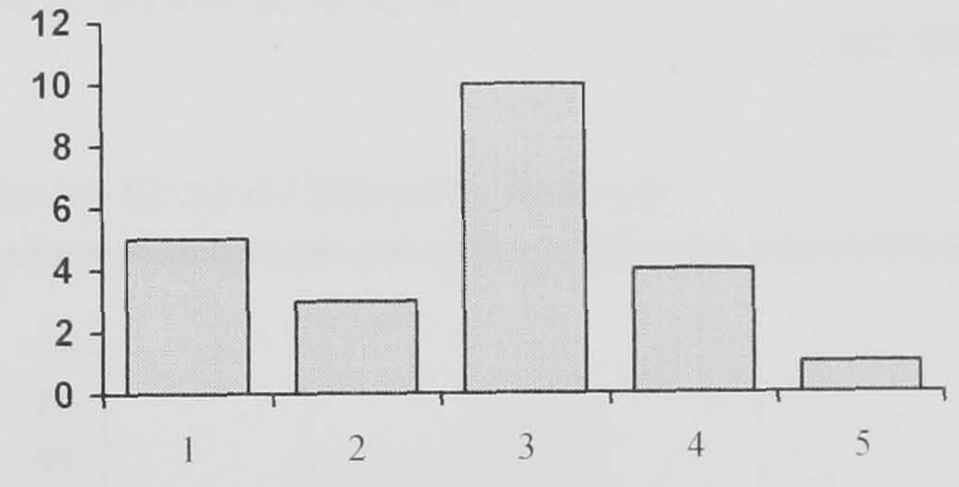
Scale: 1 useless to 5 very useful
N=22 Mean=3.10 Only Scottish secondary

Survey 49: % Pupils gaining NC modules and CSYS



Scale: 1 useless to 5 very useful
N=22 Mean=2.86 Only Scottish secondary

Survey 50: % Staying on rates



Scale: 1 useless to 5 very useful
N=23 Mean=2.70 Only Scottish secondary

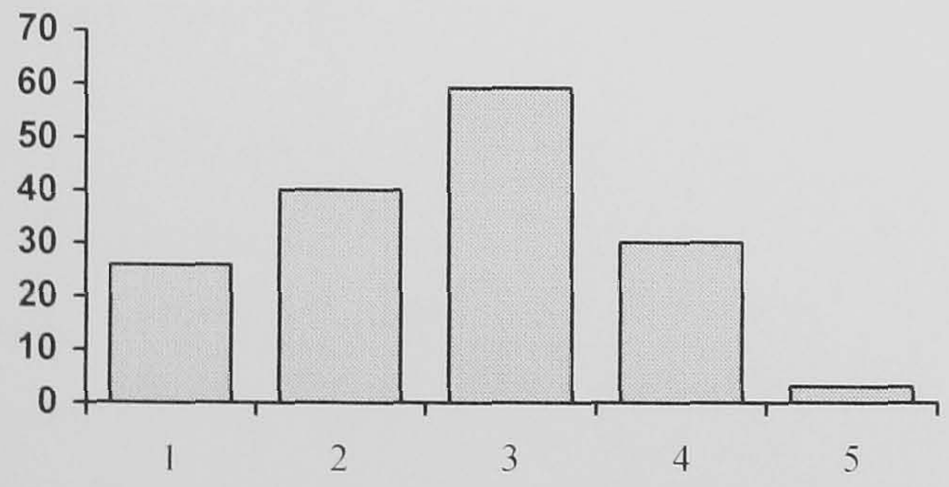
How well do KPIs indicate performance at different levels

As well as showing how well an individual school is performing, KPIs may be used to judge other levels of the system; teachers, education authorities as well as the government. The findings from *Survey 51* to *Survey 54* show that the respondents found them moderately useful, (to similar degrees), for national, (L)EA and school performance. However, they were far less sanguine about their use for judging teacher performance. Fortunately, the current threshold system (DfES 2001c) uses a range of indicators and judgements, and this allows assessors to take into account various contextual factors in terms of a teacher's results.

KPIs may show that an LEA is performing well, however underlying this may be serious implications for some schools and the education they offer. This is illustrated below with the case of a non-selective school in a 'high' performing LEA;

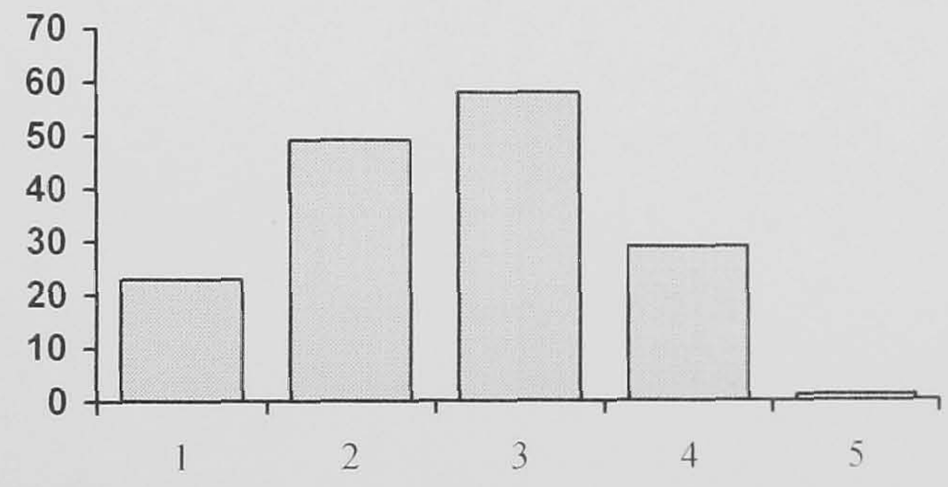
> The catchment area of my school includes 6 Grammar schools and a City Technology College. Given this level of 'creaming off' of the best students how can a league table of this area give an accurate indication of the quality of education in our school. Our school offers our pupils a vast array of educational experiences – performing arts, sport, academic and cultural links with schools from several European countries, residential field courses, outward bound courses, D of E, vocational experience / qualifications etc – that are not reflected in anyway in the KPIs.
>
> (RS 70)

Survey 51: National



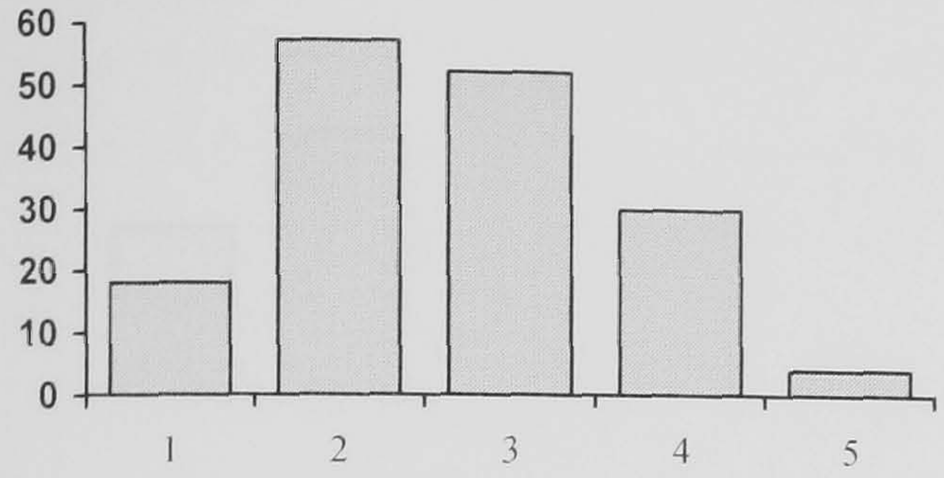
Scale: 1 useless to 5 very useful
N=158 Mean=2.65

Survey 52: LEA / Education Authority



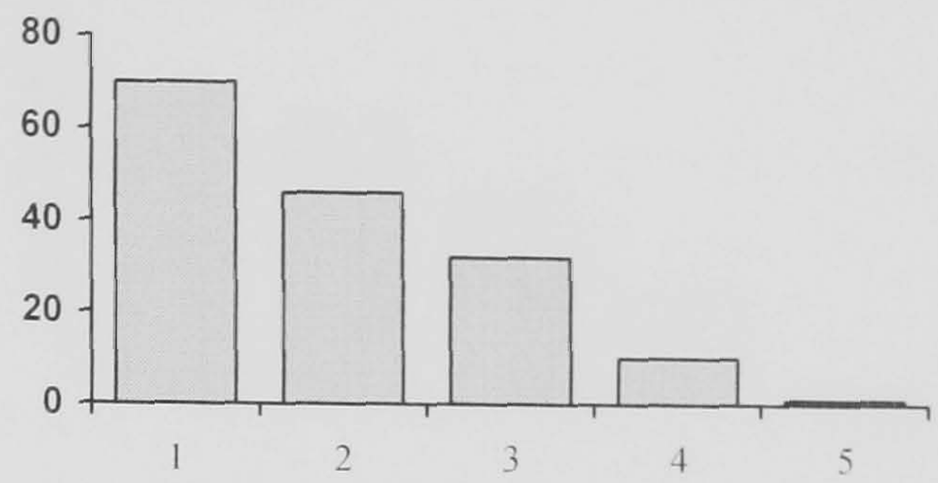
Scale: 1 useless to 5 very useful
N=160 Mean=2.60

Survey 53: School



Scale: 1 useless to 5 very useful
N=161 Mean=2.66

Survey 54: Teacher



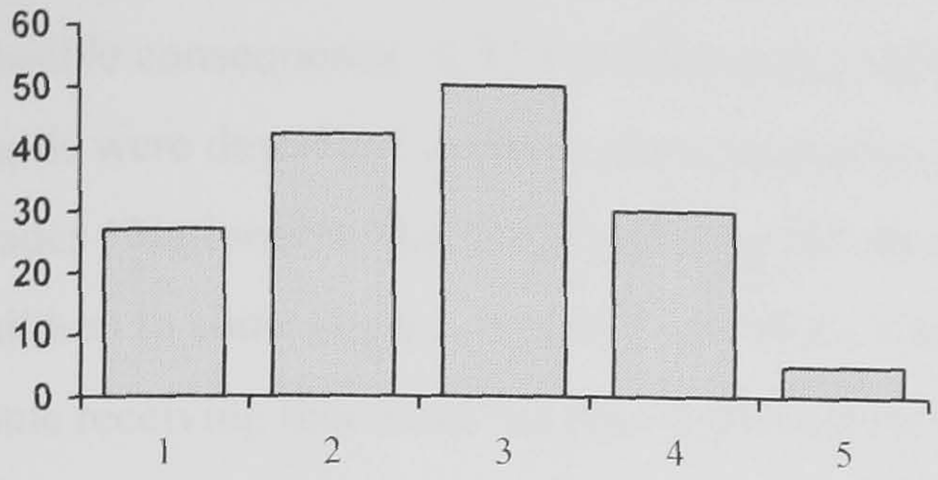
Scale: 1 useless to 5 very useful
N=159 Mean=1.91

How well do KPIs indicate performance for different time scales

KPIs may be used to judge performance for different points in time. *Survey 55* and *Survey 56* suggest that the respondents have most faith in their use as indicators of current and past performance, but are far less useful at judging future performance (*Survey 57*). This questions some of the main uses of KPIs, firstly as already discussed for parents identifying the most suitable school, and secondly for Ofsted / HMI identifying 'problem' schools; indeed as has been shown the inherent instability of KPIs[143] further questions their predictive value.

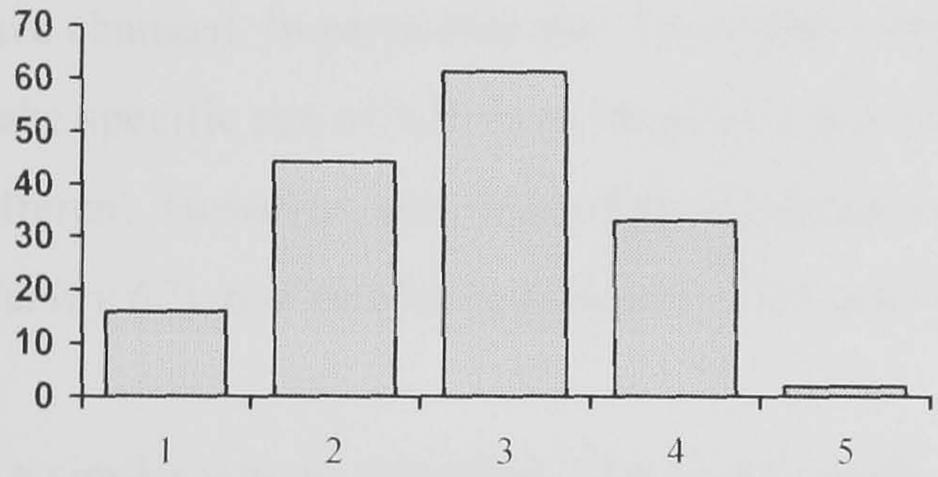
[143] Eg. due to cohort variation

Survey 55: The past



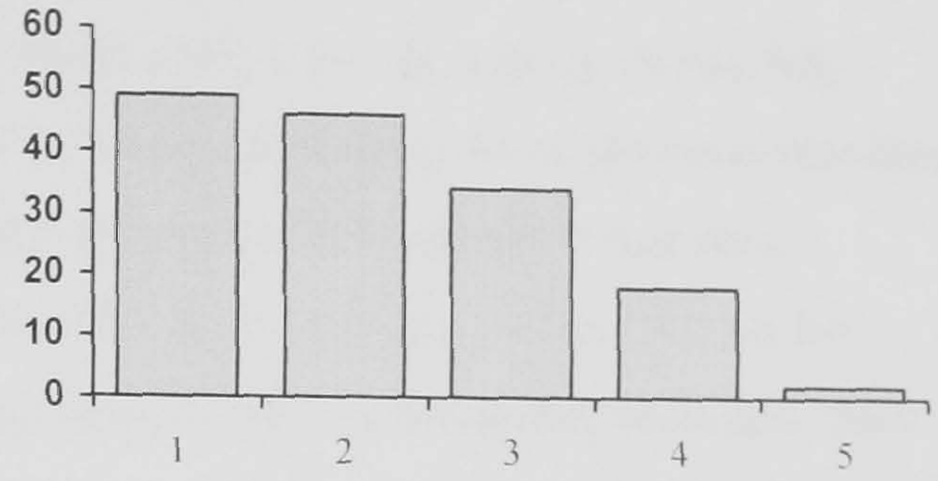
Scale: 1 useless to 5 very useful
N=154 Mean=2.64

Survey 56: Currently



Scale: 1 useless to 5 very useful
N=156 Mean=2.75

Survey 57: The Future



Scale: 1 useless to 5 very useful
N=149 Mean=2.18

KPIs and the management process

Chapter 2 discussed the use of PIs in the management of schools, and in particular the need to feedback relevant information. *Survey 58* and *Survey 59* asked whether KPI systems help identify the key strengths and weaknesses of the respondents' schools. The results were similar for both questions with very few indicating that they definitely did, but a large proportion stating that they were of some use. Their (careful) use as management tools is illustrated below;

> Personally I don't object to KPIs / indicators if they have the effect of directing my attention at an area of the curriculum that needs work – great!
>
> If I can use them to encourage (not blame) individual teachers to improve – great!
>
> (RS 204)

And in a similar way a primary school finds them useful for internal use;

> National tests are a useful tool within a school if used internally to improve / encourage learning. However to compare schools using National tests is dishonest.
>
> (RS 214)

The issue of schools concentrating resources on 'Borderline' pupils had been identified as a possible consequence of KPI systems (eg. chapter 7, Smith 1995a). For the survey 'borderline' pupils were described as those close to gaining GCSE 5 A*-C's / KS2 level 4s /credit level standard grades / National testing targets. *Survey 60* shows that a majority of the schools target these children to some degree, and with about 20% indicating that they definitely do this. Whilst for some receiving this extra support there will be obvious benefits such as being able to access other areas of the curriculum due to improves skills; for others (eg. gifted or special needs) there may be a negative impact.

Survey 61 shows that for the majority of the respondents KPIs are little used, if at all, in the appraisal process. Since the survey in 1999 the statutory requirements for appraisal in England have changed. In particular the 'Threshold' and 'Performance management' systems, both of which make specific use of KPI data; therefore it would seem likely that the responses would now[144] be different. However, very few of the respondents felt that KPIs recognised the quality of their work (*Survey 62*), and this has implications for how they should be used in appraisals.

In a similar way to appraisal, *Survey 63* suggests that KPIs do not significantly contribute to personal and professional development. A useful observation from a primary school was that less 'threatening' and low stakes indicators would be more appropriate;

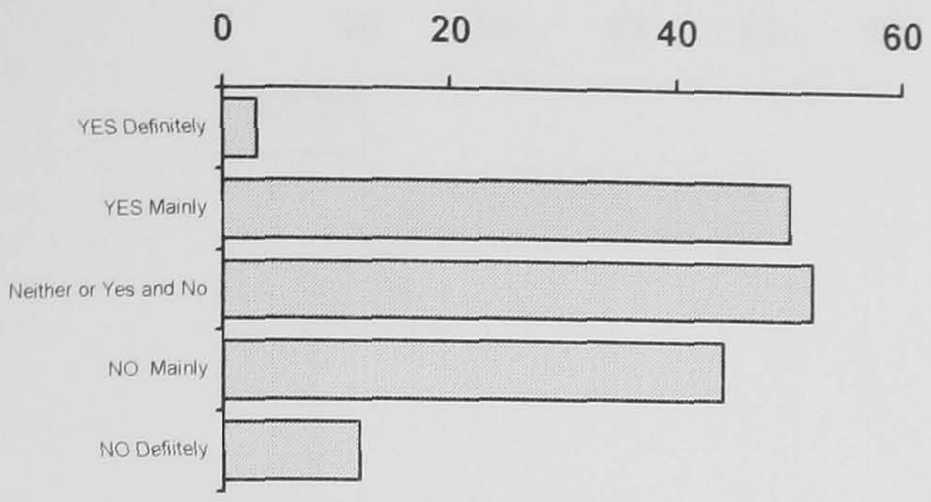
> A system of self evaluation techniques (eg Quality standards) would be more useful in involving staff in a positive way and promoting an ethos of trust and partnership.
>
> (RS 250)

The final question (*Survey 64*) in this section has already been mentioned, and shows that a high proportion of respondents do participate in their own target setting, although as discussed there are moves for these to be centrally set (Hackett 2000).

---

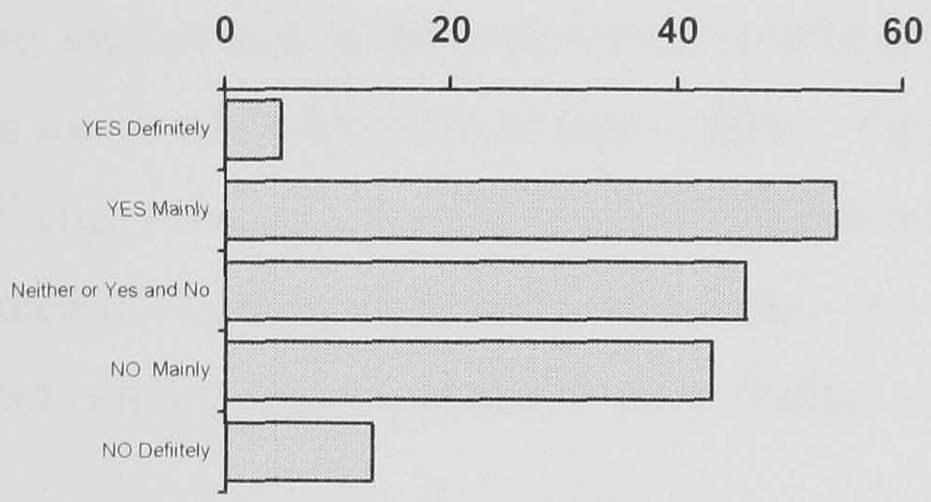
[144] 2002

Survey 58: Identify strengths



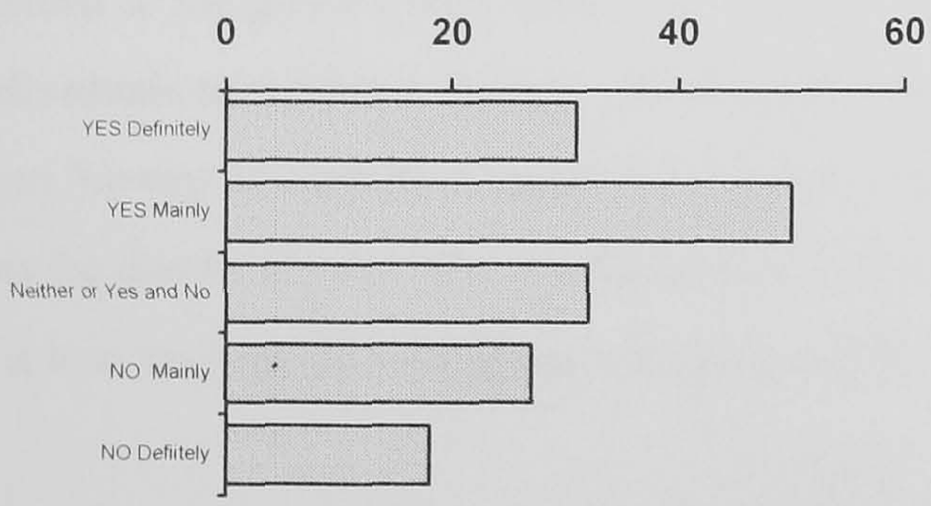
N=161 Mean=3.07

Survey 59: Identify weaknesses



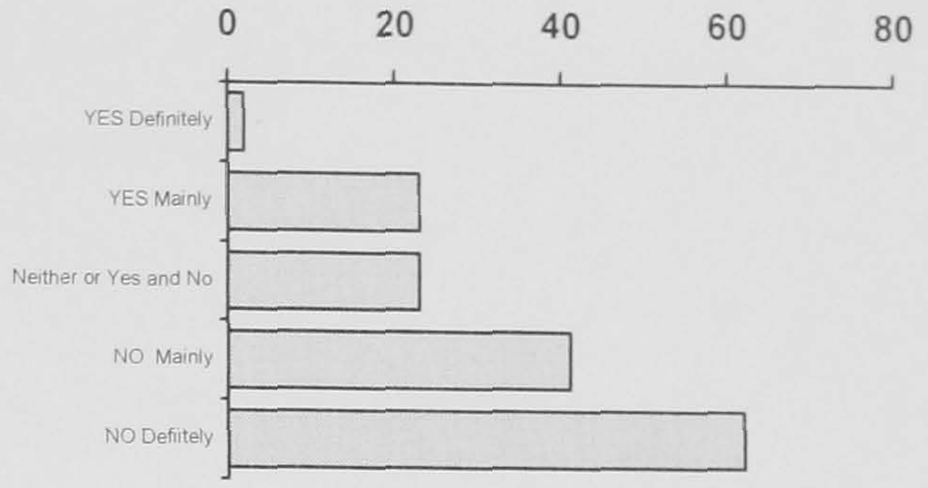
N=161 Mean=3.03

Survey 60: Target borderline pupils



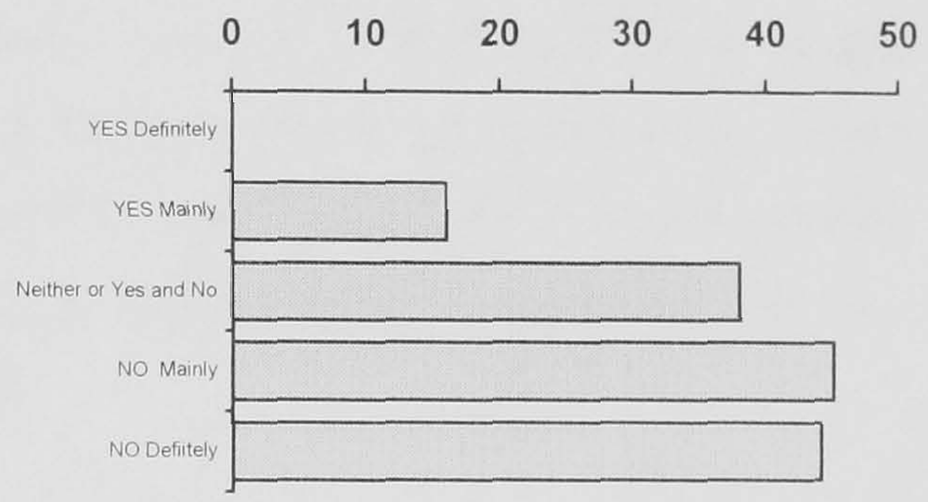
N=158 Mean=2.69

Survey 61: KPIs used for your appraisal



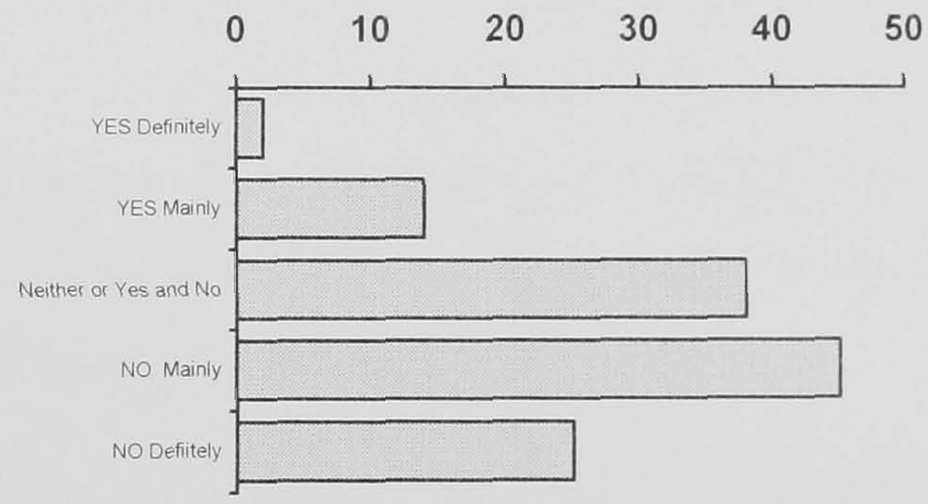
N=151 Mean=3.91

Survey 62: Recognise quality of work



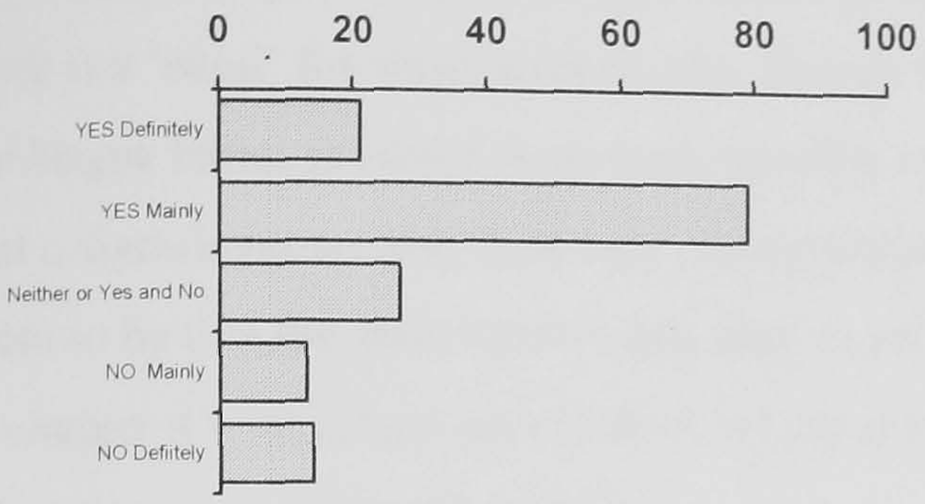
N=151 Mean=3.79

Survey 63: Contribute to professional development



N=142 Mean=3.62

Survey 64: Participate in target setting



N=154 Mean=2.48

How do you feel about publication of league tables

This section look at the respondents (not Scottish primary) views on publishing KPI data, in what are commonly referred to as league tables. Technically they are a media, not government creation, although ministers do refer to school league tables (eg. HOC 2000). In general most respondents indicated that they were bad for education (*Survey 65*), as the quote below suggests this view is most certainly not limited to 'low' performing schools;

> The public humiliation of schools does no one any good. I speak for a school which was mentioned in the Sunday Times top 50 schools in Scotland. We do very well but others are working hard and only get blame.
>
> (RS 227)

Central to the government policy of 'driving up' standards is the aim of increasing the stress on individuals to perform 'better'. Whilst some degree of stress may be good (chapter 6) the results from *Survey 66* may be a cause for concern, given the degree of the responses. Furthermore, this may be due to factors beyond the individual's control, for example pressure being exerted because of a low league table position, which in reality may be due to cohort variation.

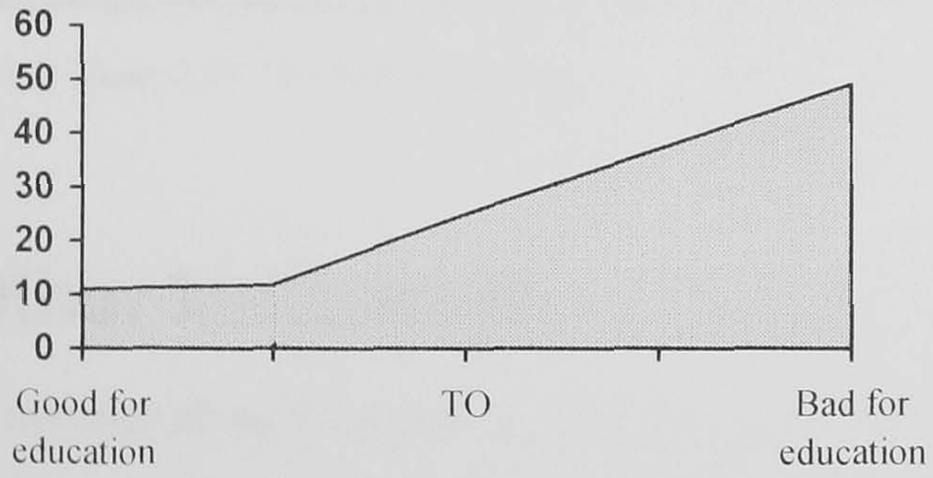
Consistent with other responses on KPIs the majority of respondents felt that league tables did not give an accurate or fair impression (*Survey 67* and *Survey 69*). For example, RS (10) points out that league tables do not reflect the problems of a very high level of turnover (both pupils and staff) for her school in Hounslow. And as previously mentioned the problems of size and cohort;

> In a small school with a large number of SEN these indicators [league tables] take NO account of special needs in an easy to read way and therefore do not pass on to the general public how much progress has been made....
>
> (RS 6)

Education by its very nature is a long-term process and the information and management systems should encourage a long-term view; although as discussed in chapter 4 the government argue that there is a 'need' for short-term results. *Survey 68* shows that the majority of respondents feel that the league tables produce short term benefits at the expense of longer term gains; the graph shows that a significant number feel very strongly that this is the case. The potential danger here would seem to be that the short-term needs may override the long-term benefits. This issue was discussed in chapter 4 with reference to MBO, which it has been argued previously is similar to the 'modern' educational management models.

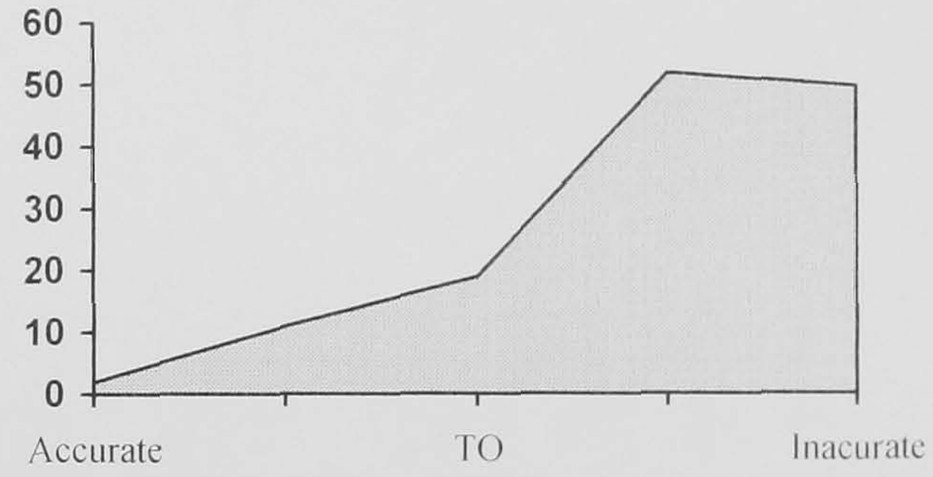
Finally, in this section, the respondents felt that league tables did not recognise quality, (*Survey 70*) and allow for local circumstances (*Survey 71*). With respect to this second point the findings are more pronounced than *Survey 24*, which looked more specifically at allowing for local circumstances in the target setting process. Given that by definition league tables are essentially local, in that they are mainly used to compare schools within an area, this is of concern, and further questions their validity.

Survey 65: Good for education



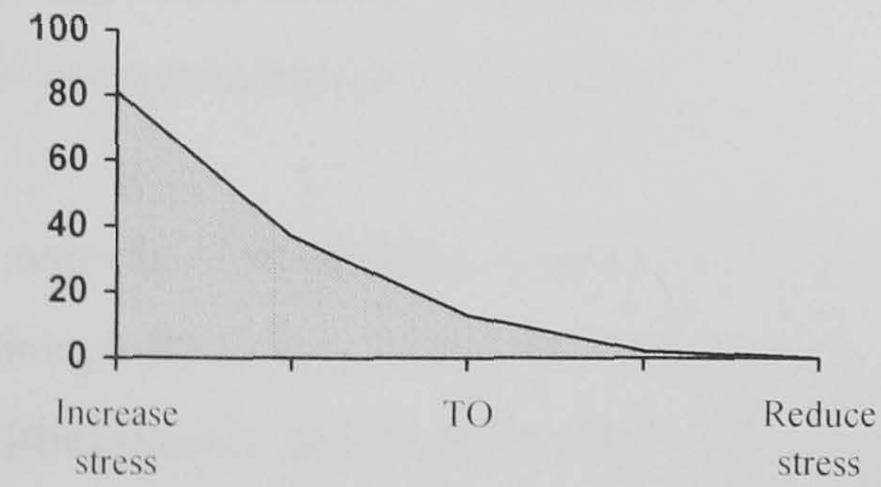
N=134 Mean=3.75 Not Scottish primary

Survey 67: Accurate impression



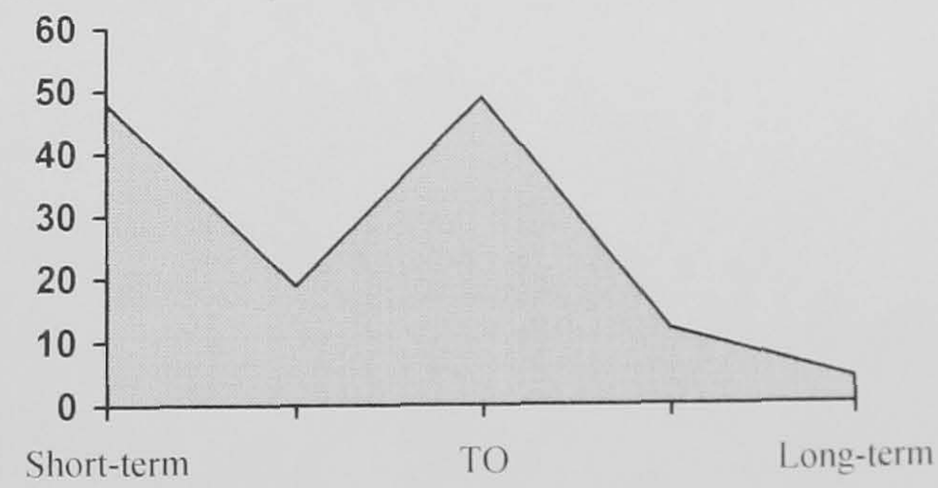
N=134 Mean=4.02 Not Scottish primary

Survey 66: Increase stress



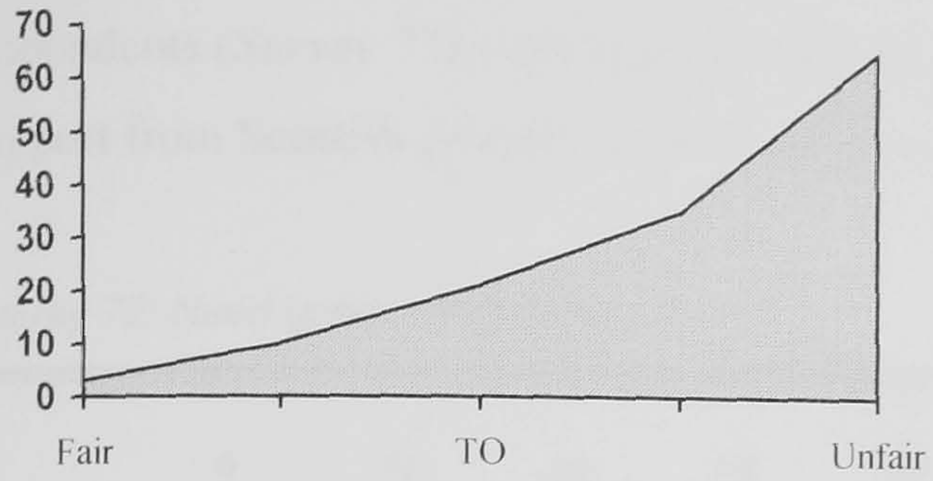
N=133 Mean=1.52 Not Scottish primary

Survey 68: Short-term gains



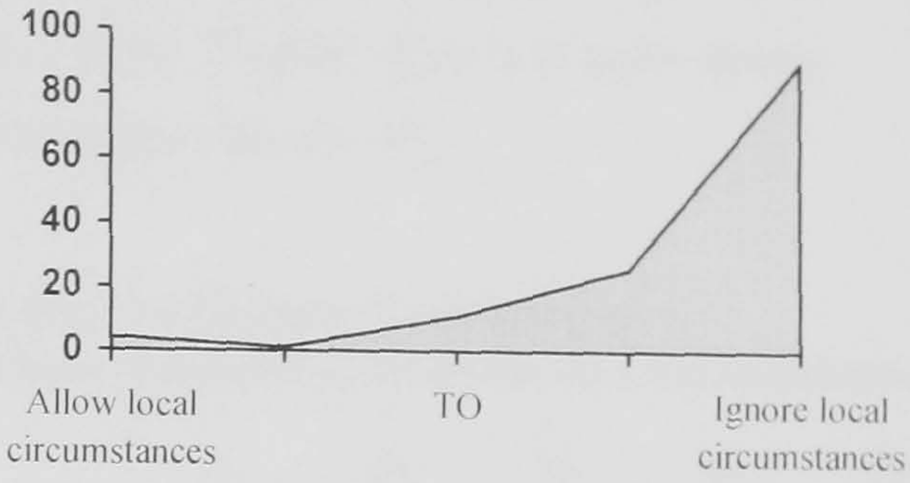
N=132 Mean=2.28 Not Scottish primary

Survey 69: Fair



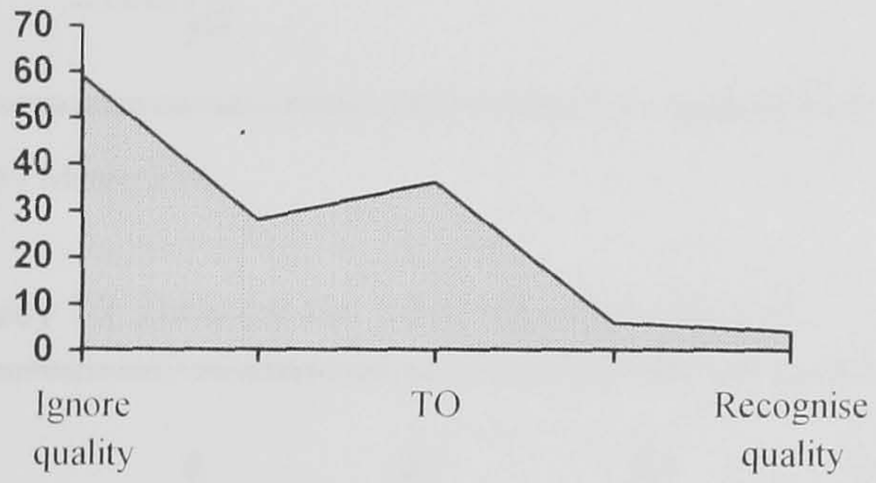
N=134 Mean=4.11 Not Scottish primary

Survey 71: Allow for local circumstances



N=133 Mean=4.50 Not Scottish primary

Survey 70: Ignore quality



N=133 Mean=2.01 Not Scottish primary

Other Indicators

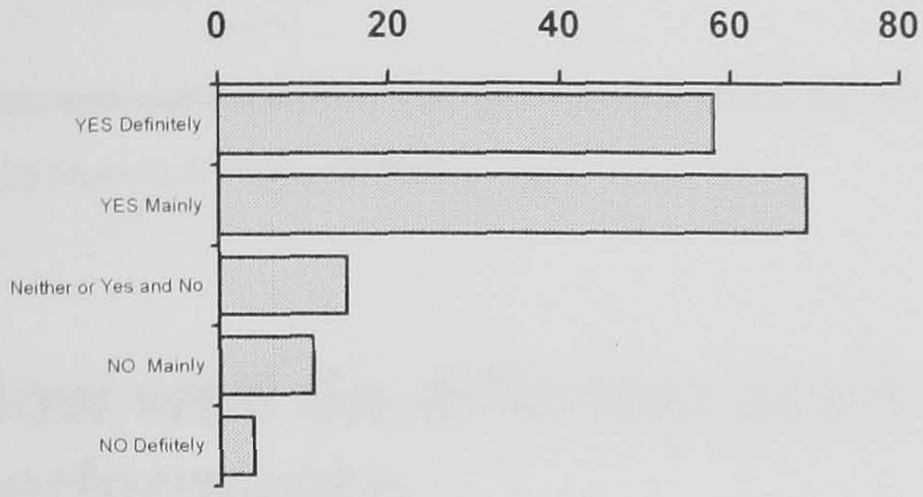
A number of more general questions were asked about possible developments in the indicator system. Firstly, respondents were asked whether schools need good numerical objective performance data, (which can complement more subjective data, such as Ofsted / HMI reports). The overwhelming majority (*Survey 72*) felt that schools did need such information, furthermore, *Survey 73* shows that the majority take the view that it is realistic to produce this data. Taken together these findings strongly support the notion that schools are not inherently against collecting and using performance information.

A majority of respondents claimed to understand the principles of 'value added' (Survey 74), although RS (23) (a maths teacher) pointed out that there was a; "high degree of misinterpretation in [the] figures, and staff need more training". Nevertheless most respondents were in favour of such indicators being included in the performance tables (*Survey 75*)[145]. These findings further support the view that schools want to use performance information, although not surprisingly most

[145] See points already discussed in this chapter for Survey 45

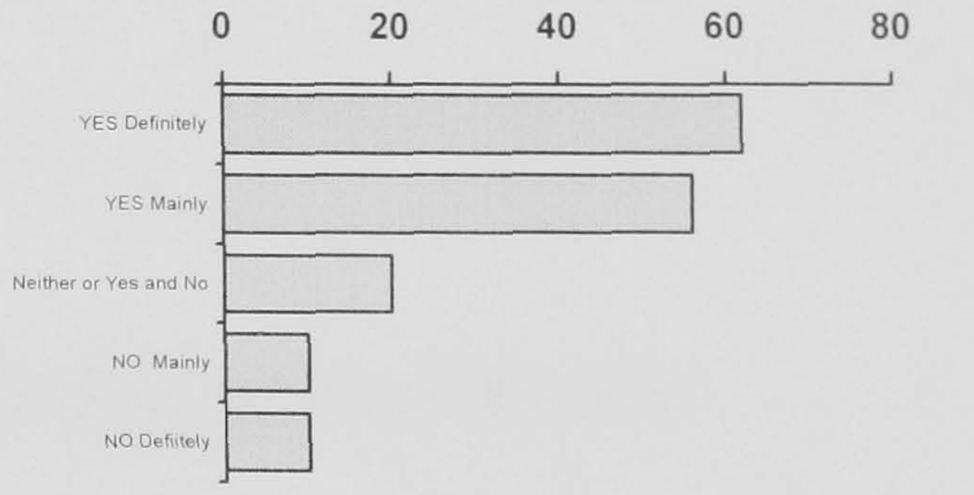
feel that like should be compared with like (*Survey 76*). There is good evidence that the performance of departments within schools vary (Fitz-Gibbon 1996), however the majority of respondents (*Survey 77*) were against making such data public. Finally, there was quite strong support from Scottish primary schools for baseline assessment (*Survey 78*).

Survey 72: Need good performance data



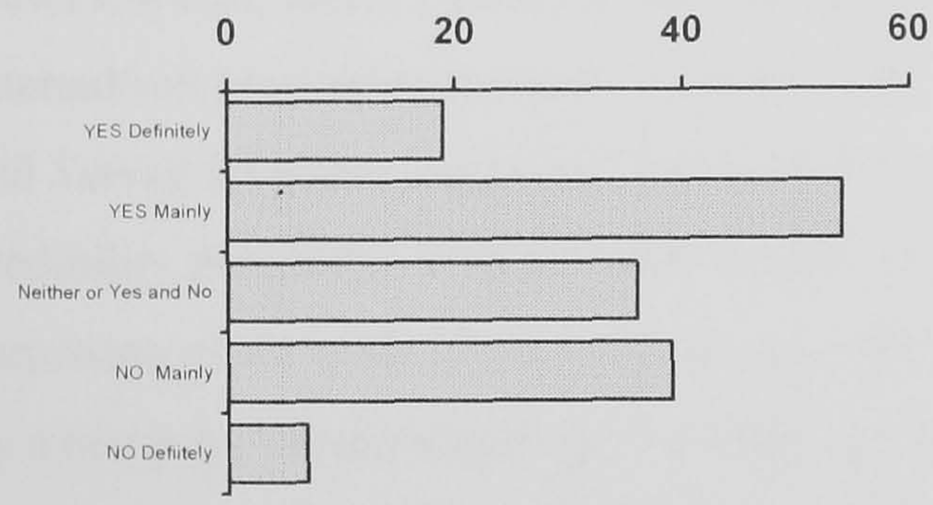
N=157 Mean=1.94

Survey 75: Indication of value added



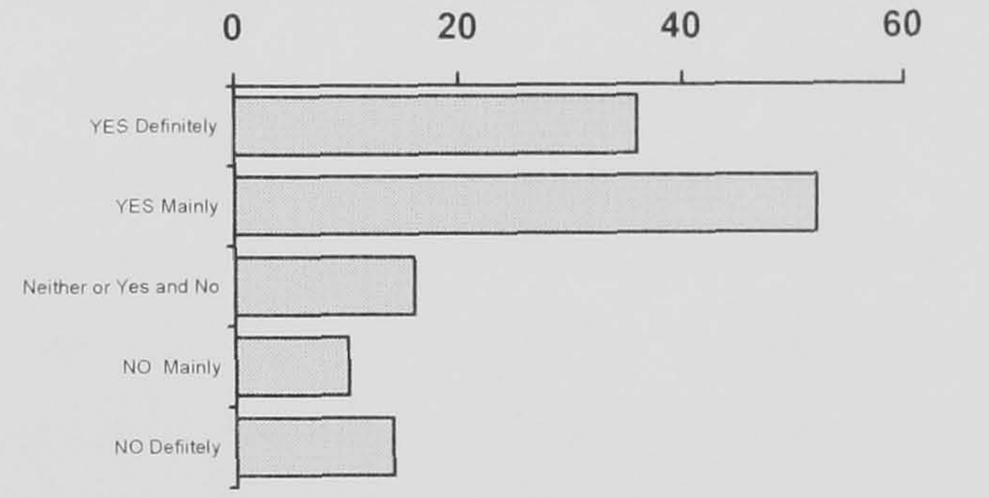
N=158 Mean=2.05 Not Scottish primary

Survey 73: Data for the most important factors



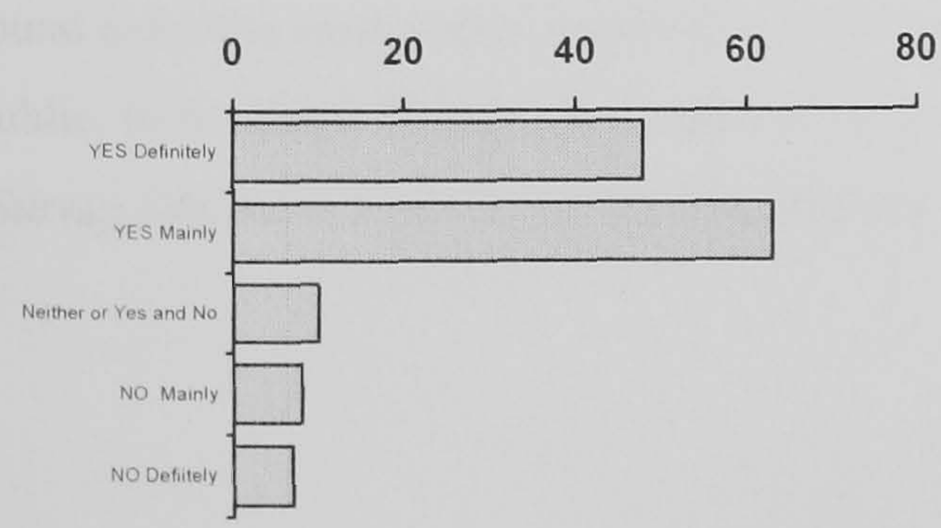
N=155 Mean=2.75

Survey 76: Compare schools similar circumstances



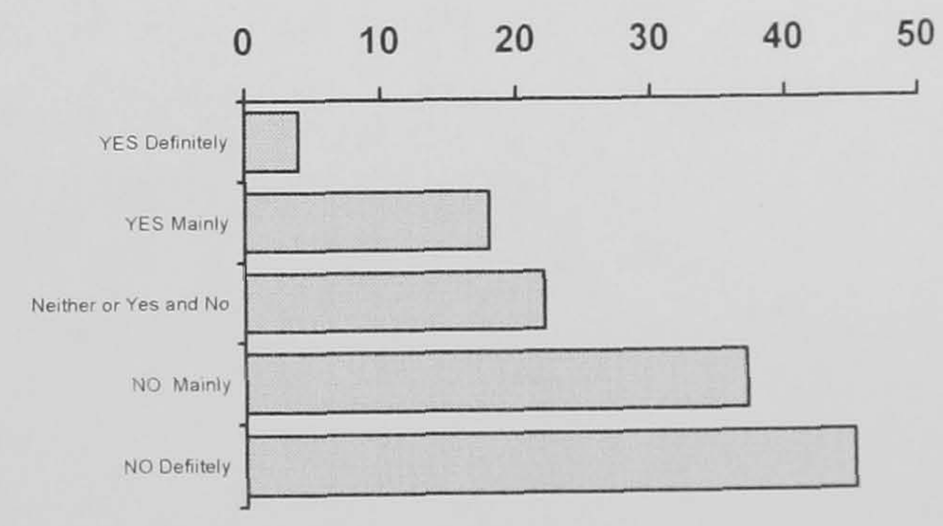
N=128 Mean=2.33 Not Scottish primary

Survey 74: Understanding of value added



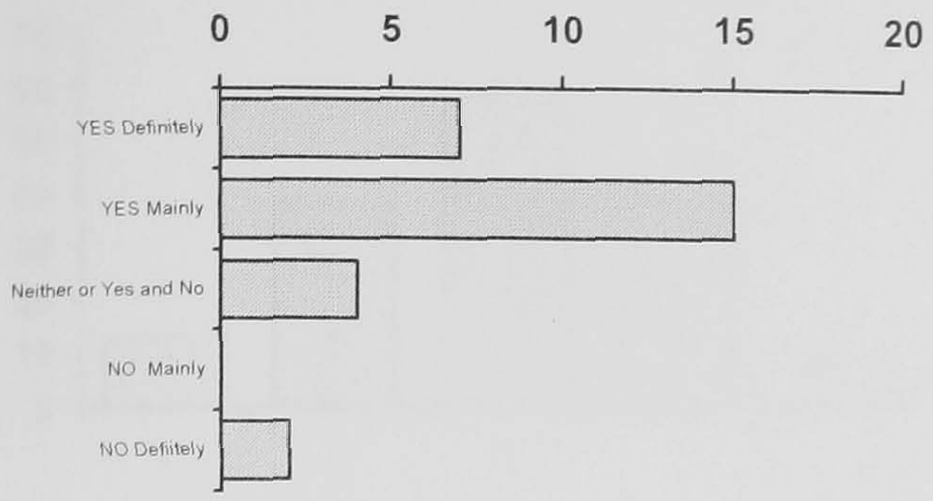
N=136 Mean=1.99

Survey 77: Department performance made public



N=126 Mean=3.80 Not Scottish primary

Survey 78: Baseline assessments



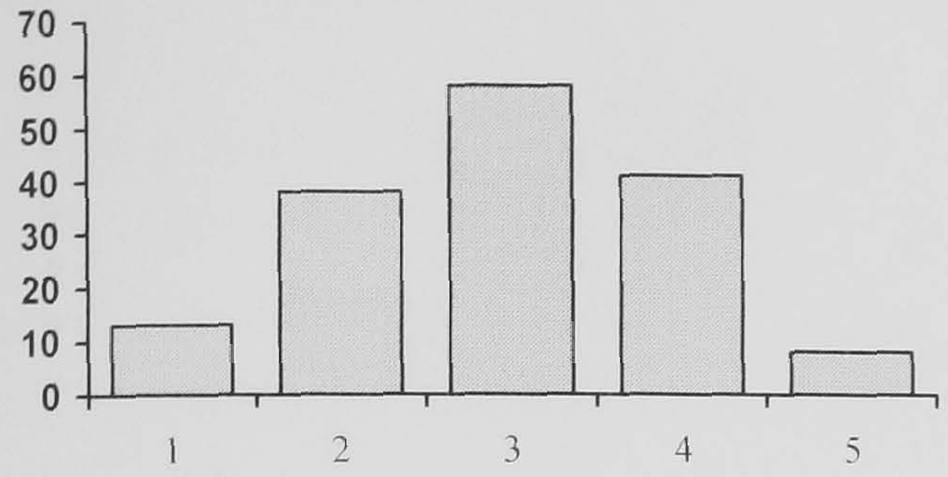
N=28 Mean=2.11 Only Scottish primary

How well do different systems judge your school's performance

This final section looks at how well the different indicator systems judge performance. Whilst the first three surveys combine data from all of the schools, due to the differences the other surveys are reported in terms of country / phase. *Survey 79*, *Survey 80* and *Survey 81*, detail the respondents views towards, KPIs, Ofsted / HMI inspections and internal /informal indicators. The latter, internal/informal systems were found to be the most useful, and this very much supports, *Survey 37* and *Survey 38* which suggested that professional judgements and internal indicators have higher credibility than KPIs. Ofsted / HMI reports were also found to be better than KPIs, which might be surprising given some of the controversy surrounding Ofsted in particular. This may be explained by a number of factors such as; the ability of inspectors being able to contextualise data, very few people reading the whole report and schools (and often the local media) highlighting / reporting the best aspects, and also the relatively small number of schools deemed to be failing.

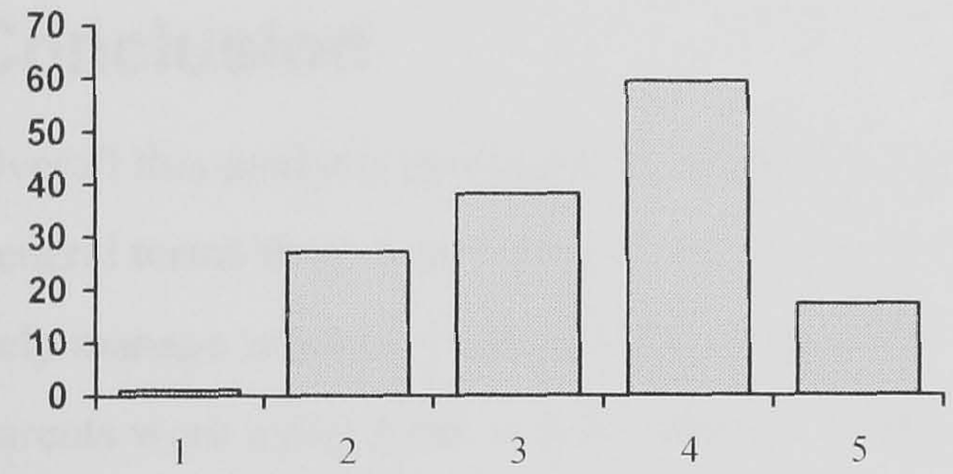
With respect to English schools, PANDA and Benchmark data (*Survey 82* and *Survey 83*) were found to be the least useful indicators. This was a little surprising given that this data is not made public. In Scotland though, the relative ratings (*Survey 85*) and to a lesser extent value added (*Survey 84*), were found to be significantly more useful than the published KPIs.

Survey 79: KPIs



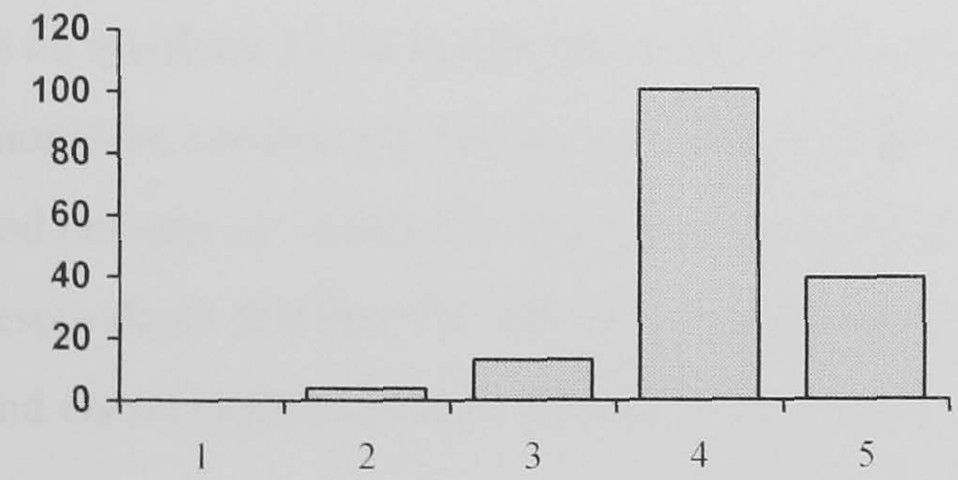
Scale: 1 useless to 5 very useful
N=158 Mean=2.96

Survey 80: Ofsted / HMI



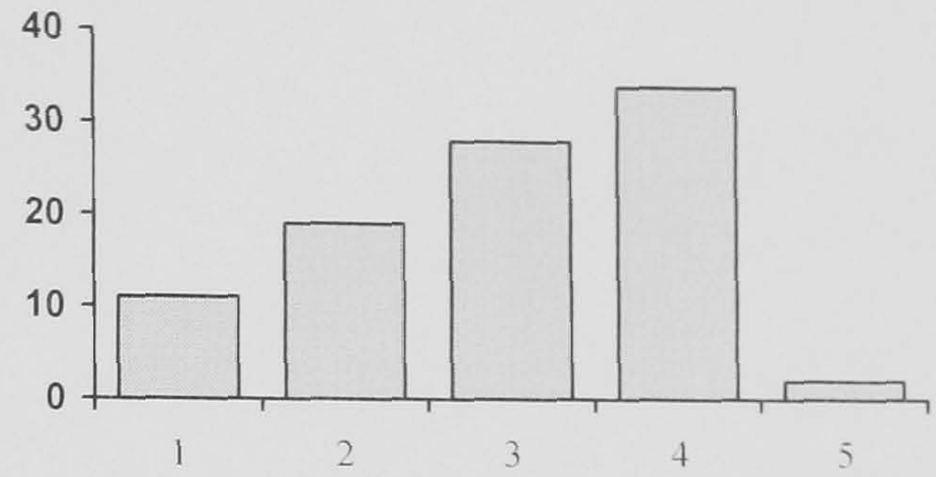
Scale: 1 useless to 5 very useful
N=142 Mean=3.45

Survey 81: Internal / Informal systems



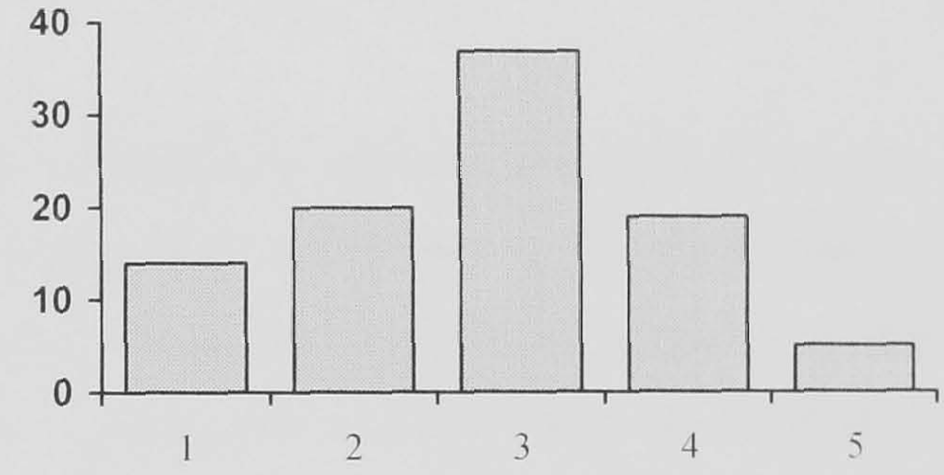
Scale: 1 useless to 5 very useful
N=156 Mean=4.11

Survey 82: PANDA



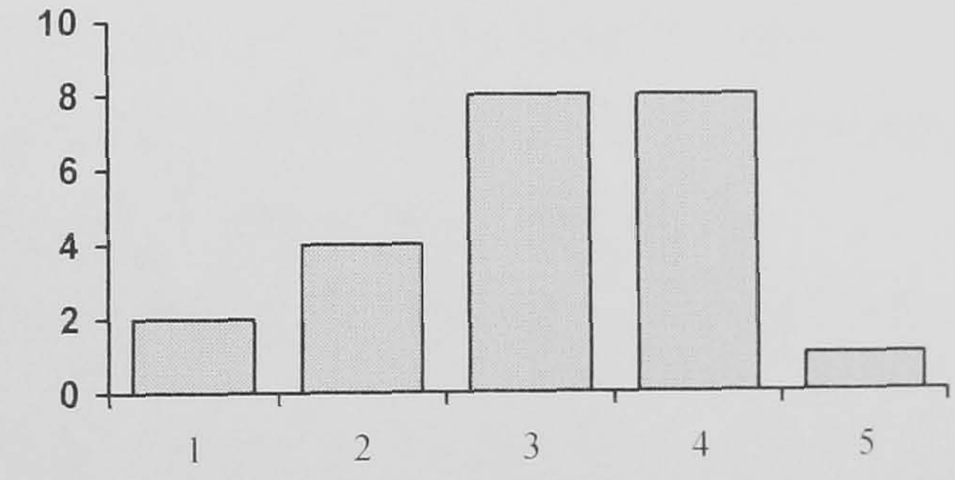
Scale: 1 useless to 5 very useful
N=94 Mean=2.97 Only English schools

Survey 83: Benchmark data



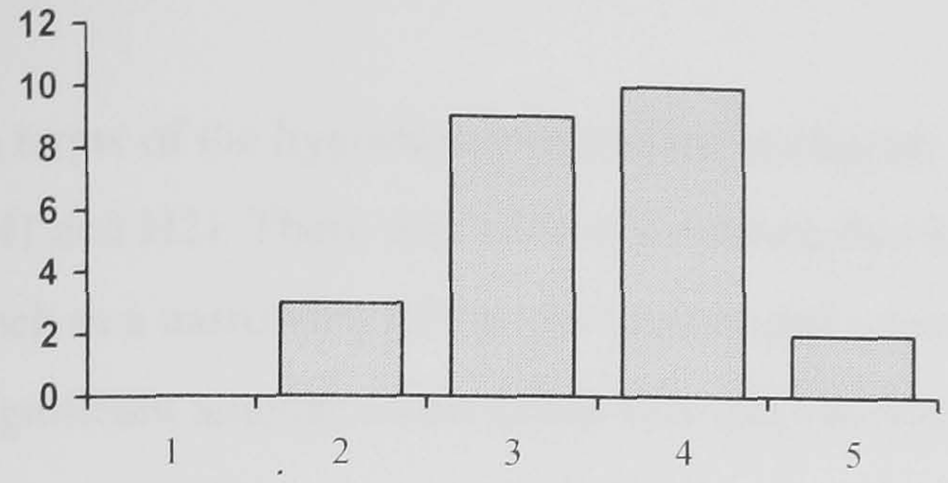
Scale: 1 useless to 5 very useful
N=95 Mean=2.80 Only English schools

Survey 84: Value Added



Scale: 1 useless to 5 very useful
N=23 Mean=3.09 Only Scottish secondary

Survey 85: Relative Ratings



Scale: 1 useless to 5 very useful
N=24 Mean=3.45 Only Scottish secondary

Conclusion

Overall this analysis produced some mixed messages and relatively few unequivocal findings. In general terms there was a good deal of support for PI systems as well as the need for good data to help manage schools. However, there was little indication, from the respondents' perspective, that parents were using KPIs to hold schools accountable. As expected there was significant evidence that KPI systems have a narrowing effect on the curriculum, and a worrying sign that many schools did not trust the systems and found them undermining. For quite a number, KPIs increase, rather than devolve 'central' control.

As would be expected most respondents felt under pressure to meet their targets, with most wanting to be involved in the setting process, as well as feeling that more aspects of the school process should be considered. However, there was evidence of targets simply being imposed from 'above', and perhaps an indication that this may be an increasing trend. Furthermore, many of the respondents felt that the indicator systems did not take account of their particular circumstances, and encouraged a 'blame' culture to develop.

League tables, not too surprisingly were felt by many to be unfair, with the media not acknowledging the limitations and many parents not being able to properly interpret and contextualise the data. In both Scottish and English secondary schools the indicators which signified higher performance (eg. 5 A*-C) were considered more useful. In general KPIs were considered more useful as indicators of current and past performance, rather than future performance. There were also signs that many of the respondents were in favour of using value added systems, although there were concerns about publishing these. In general schools felt their

'own' internal indicator systems[146] were most useful followed by Ofsted / HMI reports, and then KPIs.

In terms of the hypotheses identified in chapter 9, this analysis is relevant to the first and second (H1 and H2). There was a lot of evidence that KPI systems encourage dysfunctional behaviour, such as a narrowing of the curriculum and encouraging a 'blame culture'. Secondly, there was a significant amount of evidence that schools want, and use performance data.

In spite of the general findings it was evident, to an extent from the quantitative data and more so from the additional comments, that for many schools there was no simple one best way of using performance information; or that straightforward causes and effects could necessarily be attributed to indicator systems. Schools are complex organisations and many different factors and nuances contribute to their effectiveness, as well as 'their' reaction to external influences such as KPIs; indeed 'their' reaction to such influences may appear contradictory. This can be illustrated with comments from two of the heads taking part in the survey, one is from an English grammar school, and the other from a Scottish primary. It is worth noting that both have an 'elder statesmen' status, in that they have taught for many years and experienced the many changes which have occurred over the last thirty or so years;

> Whatever else KPIs are doing, they are changing the culture. Exam success is now 'cooler'. The cartoon academic, bald with gown, is not so much a figure of fun. In 1962 (when I started) no one cared about a class's results, or a teacher's results, or a school's results. Indeed few staff bothered to check. Now results do matter. (written by retiring bald headmaster in gown!)
>
> (RS 141)

> ....you will gather that I have a very negative attitude to national testing. I have seen many changes since I became involved with education in 1967 and I am not happy with current trends which I see as restrictive...My philosophy is aimed 100% to the individual child. National testing in maths and language is of no use whatsoever; and unnecessary in the case of skilled teachers. When National testing was imposed upon us in 1991 I fervently hoped parents would keep children off school. They cooperated 100% No P4 or P7 child attended school that day.
>
> (RS 232)

On the face of it these divergent views have little in common, however underlying these is perhaps a similar belief in what is most important.

---

[146] This includes LEA systems as national systems such as CEM indicators

Chapter 12

Comparative analyses

A great many people think they are thinking when they are merely rearranging their prejudices

(James)

Introduction

This chapter compares the results from the different countries, phases and posts (heads teachers[147]), and with particular reference to hypotheses (H3, H4, H5) detailed in chapter 9:

> **H3:** Secondary and primary schools have similar views on the use of performance data for school management, but secondaries find KPIs more dysfunctional
> **H4:** English and Scottish schools have similar views on the use of performance data for school management, but English schools find KPIs more dysfunctional
> **H5:** Teachers and heads have similar views on the use of performance data for school management, but teachers find KPIs more dysfunctional

However, whilst carrying out the analyses in this chapter it became evident that these hypotheses on their own were not very satisfactory; therefore two supplementary hypotheses (SH1 and SH2) were created.

For the analyses in this section it was decided not to compare all of the items from the questionnaire, but rather to concentrate on a number of the key issues and questions. To achieve this what were felt to be the most pertinent items were organised in to two groups: *KPIs Management* and *KPIs - Dysfunction*. Means and Standard deviations are reported, and Independent sample t-tests have been used to calculate the statistical significance (p value). A p value of less than 0.05 is commonly accepted as showing statistical significance. The significance is also shown in terms of the effect size (ES) (difference in means / pooled S.D.), and results greater than 0.80 can be considered to represent a large ES (Cohen 1977).

The selection of survey items

The first set of questions **KPIs – Management** (*Table 10*) included a range of questions from the main survey which related to the use of indicators relevant to the management of schools and the education process. They mainly looked at how the respondents use performance data.

Table 10: KPIs - Management

NAME	QUESTION	MEASURE	SURVEY Ref
Numeric Data	Do schools need good numerical objective performance data?	Likert	72
Targets Useful	Your feelings on Targets and target setting **Useful to Hindrance**	S D	27

[147] The responses from English governors are not used in this chapter

Parents	Do parents frequently refer to KPI results?	Likert	2
Pressure	Is there great pressure on your school to meet KPI Targets?	Likert	13
Participate Targets	Should you participate in the setting of your targets?	Likert	19
Judge Performance	Are the KPI results a good judge your school's performance?	1 – 5 score	79
Professional Judgements	Does KS2 / 5 – 14 testing / external examinations have higher credibility in your school than professional judgements?	Likert	38

Where appropriate the questions have been rephrased and the scales reversed to maintain meaning and context

The second set, **KPIs – Dysfunction** (*Table 11*) comprises of questions which relate to how KPIs may have dysfunctional effects. The selection aims to consider a number of different perspectives, ranging from a broad overall organisational view, through school and class perspectives, and 'down' to individual perceptions.

Table 11: KPIs - Dysfunction

NAME	QUESTION	MEASURE	SURVEY Ref
Harmful Overall	How do feel overall about KPIs in your school **Harmful to Beneficial**	S D	6
Conflicting Aims	Is there any conflict between your school's aims and its targets?	Likert	15
Concentrate Targets	Does your school concentrate on targets at the expense of other important objectives?	Likert	22
Narrow Curriculum	What effect do KPIs have on the Curriculum **Narrowing to Broadening**	S D	8
Borderline	Does your school specifically target 'borderline' pupils?	Likert	60
Blame Culture	What effect do targets have on 'the' blame culture in your school **Encourage to Discourage**	S D	29
Undermining	Overall do you personally find KPIs **Undermining to Supportive**	S D	9

Where appropriate the questions have been rephrased and the scales reversed to maintain meaning and context

For this second set in particular, the inclusion or exclusion of questions is potentially contentious, as the actual choice will affect the results. To check that these questions formed a reasonably valid basis for the analyses, a number of scenarios were run with different items and these did not

significantly affect the results. For example, thirteen questions[148] were identified as being potentially usable as indicators of dysfunction, (from which the seven were chosen). Running the various analyses with all of these gave quite similar results and levels of significance. Each of the hypotheses are assessed below in two stages, firstly with the KPIs – Management and then the KPIs – Dysfunction.

H3 Secondary and Primary Schools

This hypothesis looks at the differences between all of the secondary and primary schools in England and Scotland (Heads and Teachers). It was expected that the respondents (*Table 12*) would have similar views on the use of performance data in the organisation and management of schools.

Table 12: KPIs – Management: Primary and Secondary

QUESTION	MEAN		STANDARD DEVIATION		SIGNIFICANCE	
	Secondary	Primary	Secondary	Primary	t-Test p value	Effect size
Numeric Data	1.58	2.48	0.67	1.13	.000	1.04
Targets Useful	2.39	2.73	1.15	1.10	.072	0.30
Parents	3.78	4.11	1.17	0.95	.073	0.31
Pressure Targets	1.90	1.90	0.77	0.86	.989	0.00
Participate Targets	1.28	1.40	0.50	0.74	.231	0.20
Judge Performance	3.05	3.12	1.07	0.99	.698	0.07
Professional Judgements	3.71	3.58	0.83	1.06	.421	-0.14

Mean: A score of 1.00 would represent all respondents strongly AGREEING with the question or the first term in the semantic differential
A score of 5.00 would represent all respondents strongly DISAGREEING with the question or the second term of the S.D.
Secondary N = 79 Primary = 67 (England and Scotland)

In general terms, with the notable exception of the perceived need for data, there are few significant differences in these responses. Each of the questions is discussed below in more detail:

Numeric Data: As reported in *Survey 79* there was very strong support for the use of good quality numeric data, however these results suggest that secondary schools have a significantly greater need for performance data. This may well be explained by differences in size and organisation. Secondary schools are generally larger, and therefore more sophisticated information systems may well be considered necessary. They are more likely to have computerised administration systems which inherently 'encourage' the collection of data. Primary schools are generally organised on the

[148] Survey questions: 6,7,8,9,10,15,22,27,28,29,31,32,33

basis of a single class teacher, and this teacher may well have less need for detailed performance data in their day to day work.

Targets Useful: Both primary and secondary schools would appear to quite similar degrees, to be broadly in favour of targets, indeed for all schools it is now a legal requirement. It had been expected that for the reasons alluded to above, secondary schools would be more in favour.

Parents: It had been expected that parents of secondary school children would be more likely to make specific reference to schools about KPIs, because for secondary age children KPIs in the form of GCSEs / Standard grades are more important (to the child although perhaps not the school).

Pressure Targets: This was a quite a surprising result, it had been expected that secondary schools would be under greater pressure to meet targets. Their league tables tend to be given greater prominence in the media, and in theory at least, they are competing for students. At the primary level there is less competition, with children generally going to 'their' local school.

Careful consideration was given about the inclusion of this item, as there are arguments that it should have been included in KPIs – Dysfunction. Clearly too much pressure will be dysfunctional, however the principle of increasing pressure on schools to perform 'better' is very much at the heart of 'new educational management'. Therefore it is included, as a feature of school management, rather than a sign of dysfunction.

Participate targets: Not surprisingly all schools were keen to be involved in their own target setting, again supporting the notion that Heads and Teachers are interested in a positive way both in target setting and the use of performance indicators.

Judge Performance: Overall there was very close agreement on this question with respondents being largely neutral about the value of KPIs for judging performance. The question was asked within the context of other indicators and in both England and Scotland KPIs were considered the third best at judging performance, after internal systems, (such as Durham CEM), followed by Ofsted / HMI inspections. In England Panda and then Benchmark data were considered the poorest judges of performance, although in Scotland relative ratings and value added from *How Good Are Our Results* were considered more useful.

Professional Judgements: In a similar way there was close agreement that KPIs did not have significantly greater credibility than professional judgements. On the face of it an encouraging

view, if it is that heads and teachers are able to take this position. Although one suspected that many impassioned discussions on this principle will have taken place with others such as the governors, L(EA)s and Ofsted.

Overall these findings do essentially support the first part of the hypotheses, in that primary and secondary schools do broadly appear to have similar needs for performance information. The second part, relating to potential dysfunctional effects is considered below and it was expected that secondary schools would find KPI systems more dysfunctional.

Table 13: KPIs – Dysfunction: Secondary and Primary

QUESTION	MEAN		STANDARD DEVIATION		SIGNIFICANCE	
	Secondary	Primary	Secondary	Primary	t-Test p value	Effect size
Harmful Overall	3.28	3.03	1.28	1.30	.242	-0.19
Conflicting Aims	3.63	3.15	1.04	1.10	.007	-0.45
Concentrate Targets	3.13	2.84	1.17	1.19	.160	-0.25
Narrow Curriculum	2.19	1.65	1.13	0.87	.002	-0.53
Borderline Children	2.29	3.12	1.13	1.30	.000	0.69
Blame Culture	2.25	1.91	0.95	.094	.035	-0.58
Undermining	2.46	2.52	1.26	1.25	.771	-0.05

Mean: A score of 1.00 would represent all respondents strongly AGREEING with the question or the first term in the semantic differentials
A score of 5.00 would represent all respondents strongly DISAGREEING with the question or the second term of the S.D.
Secondary N = 79 Primary = 67 (England and Scotland)

These results quite clearly show that this second part of the hypothesis is not supported by the findings, with the primary schools appearing to report that their KPIs are more dysfunctional. The individual questions are discussed below.

Harmful Overall: Very much in support of *Survey 6*, the respondents were largely neutral in terms of their views as to whether KPIs were harmful or beneficial. It had been expected that secondary schools, with the greater prominence of KPIs, would tend to find KPIs more harmful.

Conflicting Aims: Primary schools found greater conflict between their broad aims and their KPI targets. This is not entirely surprising given that one of the main aims for primary schools is preparing children for secondary education, and for a number this might conflict with meeting specific targets. And the items measured by these targets will not be of any real use to the child when they leave the school, unlike GCSEs / Standard grades for secondary children.

Concentrate Targets: There was little evidence that either primaries or secondaries concentrated in meeting targets at the expenses of other important objectives. It had been assumed that they would be more of an issue for secondary schools, given the more obvious measures by which they are judged, and the widespread publication of their KPIs.

Narrow Curriculum: This looked more specifically at the subjects that are actually taught as part of the curriculum. There was very strong support (*Survey 8*) for the notion that KPI systems have a narrowing effect, however it was surprising to find that this was more of a problem for primary schools than secondary. This might well have longer term negative consequences for many of the children, by effectively restricting what they can learn.

Borderline Children: This considers the issues at the classroom level, and the attention which particular groups of children receive. *Survey 60* showed that the majority of schools were targeting 'borderline' children, and this analysis shows, as was expected, that this occurred more at the secondary level. There have been many claims that schools are targeting children who with extra help can, for example, increase a GCSE grade D to a C. For many receiving this extra help this is clearly not dysfunctional and may well be of benefit, however those 'comfortably' on Bs or Es may loose out. It is worth noting that the government has recently (2002) discontinued the highest level KS2 tests, and this might be viewed as an intimation to concentrate on the lower, measured (4 and 5) levels.

Blame Culture: *Survey 29* showed that this was an issue for many of the schools, however very surprisingly (and of great concern) this appear to is more of an issue for primary schools. It had been assumed that because primary school are smaller and the head is often also a teacher, there would be less of a blame culture. Such a culture can of course be very demoralising and may well have an negative impact on the overall quality of the education in its broadest sense.

Undermining: This final question considers whether respondents themselves found that KPI systems are undermining or supportive of their work. There was very little difference between primary and secondary schools, and overall limited evidence (*Survey 9*) that this was a significant problem.

As was expected the first part of the analysis showed that primary and secondary schools had similar needs, and used KPI systems in similar ways; however the second half was rather confounding. There was no evidence that overall (on the definition in this thesis) the secondary schools found their KPI systems more dysfunctional, on the contrary the findings were if anything

the other way. Therefore this hypothesis is rejected, although it would seem that more exploration and investigation of the findings is appropriate. This is carried out later in the chapter.

H4 English and Scottish Schools

This hypothesis looks at the differences between all of the English and Scottish schools (Primary and Secondary). It was expected that the respondents (*Table 14*) would have similar views on the use of performance data in the organisation and management of schools.

Table 14: KPIs – Management: England and Scotland

QUESTION	MEAN		STANDARD DEVIATION		SIGNIFICANCE	
	England	Scotland	England	Scotland	t-Test p value	Effect size
Numeric Data	1.84	2.24	0.92	1.13	.022	0.40
Targets Useful	2.55	2.54	1.17	1.07	.929	-0.01
Parents	3.88	4.02	1.05	1.14	.452	0.13
Pressure Targets	2.01	1.70	0.82	0.74	.026	-0.39
Participate Targets	1.33	1.35	0.66	0.55	.811	0.03
Judge Performance	3.11	3.03	1.03	1.05	.688	-0.08
Professional Judgements	3.67	3.61	0.90	1.02	.699	-0.06

Mean: A score of 1.00 would represent all respondents strongly AGREEING with the question or the first term in the semantic differential
A score of 5.00 would represent all respondents strongly DISAGREEING with the question or the second term of the S.D.
England N = 92 Scotland = 54 (Heads and Teachers)

In general these findings support the first part of the hypothesis, although two aspects are worthy of more detailed discussion. The analysis shows that English schools feel more strongly that they need numerical performance data. Given the higher prominence of KPIs in England, (eg. league tables for both primary and secondary schools), this may partly explain the desire for schools here to collect data which may provide a broader and more valid picture of their performance. Furthermore, 3rd party systems such as those from Durham CEM are more widespread in England and therefore schools may be more convinced of the potential benefits.

The finding that Scottish schools are under greater pressure to meet targets was surprising. Little clear evidence to explain this was found on the questionnaire, although a number of Scottish schools mentioned their relationship with their EA. This may provide a clue as to what is behind this finding. In England the government's policy has been very much to sideline LEAs from the day to day management of schools, although they do still have a responsibility for agreeing targets. However, in Scotland EAs have a closer relationship with their schools, and from this position may

have been able to put more or additional pressure on the schools to meet targets. This is a very tentative explanation, and an issue worthy of further investigation.

Overall, these findings do essentially support the first part of the hypotheses. The second part, relating to potential dysfunctional effects is considered below, and it was expected that English schools would find KPI systems more dysfunctional.

Table 15: KPIs – Dysfunction: England and Scotland

QUESTION	MEAN		STANDARD DEVIATION		SIGNIFICANCE	
	England	Scotland	England	Scotland	t-Test p value	Effect size
Harmful Overall	2.95	3.53	1.24	1.30	.007	0.46
Conflicting Aims	3.28	3.63	1.12	1.01	.064	0.33
Concentrate Targets	2.95	3.09	1.20	1.17	.470	0.12
Narrow Curriculum	1.71	2.33	0.96	1.10	.001	0.61
Borderline Children	2.35	3.23	1.16	1.29	.000	0.72
Blame Culture	1.96	2.33	0.95	0.93	.023	0.39
Undermining	2.25	2.91	1.25	1.13	.002	0.55

Mean: A score of 1.00 would represent all respondents strongly AGREEING with the question or the first term in the semantic differentials
A score of 5.00 would represent all respondents strongly DISAGREEING with the question or the second term of the S.D.
England N = 92 Scotland = 54 (Heads and Teachers)

These findings show very strong support for the second part of the hypothesis, namely that English schools would find KPI systems more dysfunctional. Indeed the only 'dissenting' items were how much the aims conflict with the targets and related to this how much schools concentrate on targets at the expense of other important objectives.

Although these findings do support this hypothesis, as with the previous one (H3) there would appear to be more to this than simply the difference between the countries (or phase). Therefore, as mentioned before this is explored further later in this chapter.

H5 Heads and Teachers

This hypothesis looks at the differences between all of the heads and teachers in England and Scotland (Primary and Secondary). It was expected that there would be some differences in the views (*Table 16*) on the use of performance data in the organisation and management of schools given the differing roles of heads and teachers.

Table 16: KPIs – Management: Heads and Teachers

QUESTION	MEAN		STANDARD DEVIATION		SIGNIFICANCE	
	Heads	Teachers	Heads	Teachers	t-Test p value	Effect size
Numeric Data	1.89	2.15	1.03	0.97	.157	0.26
Targets Useful	2.59	2.46	1.12	1.17	.510	-0.11
Parents	3.91	3.96	1.10	1.04	.852	0.05
Pressure Targets	1.80	2.11	0.78	0.83	.031	0.39
Participate Targets	1.30	1.41	1.30	1.41	.311	0.08
Judge Performance	3.18	2.87	1.05	0.97	.090	-0.31
Professional Judgements	3.65	3.65	1.00	0.82	.990	0.00

Mean: A score of 1.00 would represent all respondents strongly AGREEING with the question or the first term in the semantic differential
A score of 5.00 would represent all respondents strongly DISAGREEING with the question or the second term of the S.D.
Heads N = 100 Teachers = 46 (England and Scotland)

The overall degree of agreement was quite surprising, particularly given the increasingly managerialist role of heads (ie. NPM chapter 7). This could encourage the use or possible misuse of PIs as control mechanisms. For example, the target setting process could be seen by teachers as a means of ensuring conformity, and applying pressure. In terms of pressure to meet targets it was expected that teachers would feel under greater pressure, and furthermore these may have been set arbitrarily and without wholehearted agreement. However, the opposite was the case with heads feeling under significantly greater pressure to meet KPI targets. A possible explanation may be that heads are attempting to 'protect' staff from what may be iniquitous school targets. Given the relative 'isolation' of the heads position, this may have negative effects on them personally, and explain the very sad and increasingly frequent examples of heads 'cheating' in the KS2 tests.

It was expected that teachers would report more discussions with parents about KPIs, however this was not the case. In most situations class teachers will have far more contact with parents, and it would seem likely that parents would be more willing to talk to them about test and exam performance; indeed many parents would probably not feel happy discussing such issues with the head.

As with the first group of questions it was also expected to find some significant differences in terms of the dysfunctional effects of KPIs (*Table 17*).

Table 17: KPIs – Dysfunction: Head and Teachers

QUESTION	MEAN		STANDARD DEVIATION		SIGNIFICANCE	
	Heads	Teachers	Heads	Teachers	t-Test p value	Effect size
Harmful Overall	3.17	3.16	1.04	1.35	.951	-0.01
Conflicting Aims	3.27	3.71	1.00	0.97	.021	0.45
Concentrate Targets	3.04	2.91	1.21	1.01	.549	-0.12
Narrow Curriculum	1.89	2.07	0.55	0.92	.350	0.25
Borderline Children	2.69	2.62	1.27	1.18	.757	-0.06
Blame Culture	2.10	2.06	0.73	0.95	.850	-0.05
Undermining	2.45	2.58	1.26	1.04	.570	0.11

Mean: A score of 1.00 would represent all respondents strongly AGREEING with the question or the first term in the semantic differentials
A score of 5.00 would represent all respondents strongly DISAGREEING with the question or the second term of the S.D.
Heads N = 100 Teachers = 46 (England and Scotland)

Again there was surprisingly close agreement between heads and teachers, with the only statistically significant difference being in terms of school aims conflicting with the targets. This might be explained by heads being more aware of the longer term more strategic issues facing the school, and how these might conflict with the targets. Furthermore, heads are perhaps more likely to discuss such issues with other heads, their governors and advisers, which may make them more aware of the conflicts. It may however be that teachers are increasingly seeing their aims as meeting their targets, in which case there will of course be little conflict, although where this will take education is another matter.

There was widespread agreement that KPIs have a narrowing effect on the curriculum. It had been expected that teachers would have stronger feelings on this than heads, given their closer proximity to the children and the actual learning process. Again there was no statistical evidence to suggest that this was the case.

From an organisational and management point of view, on the face of it the results from these analyses are of course very encouraging. There can be very few other industries where 'workers' and 'managers' are in such apparent harmony. Therefore these findings do not support the hypothesis. However, great care should be taken with reading too much in to this, primarily because of potential short-comings in the sample.

The problems of response rate (chapter 10) and non-returns is potentially greater within this sample of teachers, largely because it can not be claimed with certainty that heads distributed the

questionnaires as intended. This might have led to some degree of the head selecting or self selecting, which may have resulted in a sample of teachers whose views might not be very representative of teachers in general, and more in accordance with their head. This pessimistic interpretation may of course be wrong, nevertheless further corroborative evidence should be gained before coming to any firm conclusions.

Supplementary hypotheses

Whilst carrying out the various analyses in this chapter, and to an extent before that during the initial data screening, it became apparent that the hypotheses (H3 and H4 in particular) were not entirely satisfactory, and were not considering some key issues and differences in the data. It was the items which comprised the KPIs – Dysfunction group that were of greatest concern, and in particular what appeared to be the differences between English and Scottish primaries. Therefore to help explore these issues a new scale, *Dysfunction Index* was formed by combining the seven items from KPIs Dysfunction. An analysis of internal reliability gave a Cronbach Alpha value = 0.80, suggesting quite a high level of reliability (Thietart *et al* 2001). An initial analysis of this new scale was carried out (*Table 18*):

Table 18: Dysfunction Index - Univariate analysis

	TYPE III SUM OF SQUARES	F	Sig.
Country	11.490	30.462	.000
Post	.327	.866	.354
Type	.309	.820	.367
Country * Post	1.291	3.422	.067
Country * Type	7.593	20.132	.000
Post * Type	.307	.814	.368
Country * Post * Type	.402	1.065	.304

Country = England / Scotland
Post = Heads / Teachers
Type = Secondary / Primary

This confirms the significant difference between countries, and also shows the importance of different types of schools and the country in which they are located. The next step was to calculate the means and confidence intervals (*Table 19*) for each type of school in each country, to give an indication of the degree of difference:

Table 19: Dysfunction Index - Descriptive

	N	MEAN	STANDARD DEVIATION	STANDARD ERROR	95% CONFIDENCE INTERVAL FOR MEAN	
					Lower	Upper
English Secondary	53	2.7192	.6456	.0887	2.5413	2.8972
English Primary	39	2.1757	.5711	.0914	1.9906	2.3608
Scottish Secondary	26	2.8013	.5987	.1197	2.5542	3.0484
Scottish Primary	26	3.1958	.6452	.1242	2.9405	3.4510
ALL	146	2.6756	.7048	.0587	2.5595	2.7917

Mean - Low number = more dysfunction, high number = less dysfunction

What stands out most obviously is the substantial difference between the English primaries and the others. A Post Hoc Test, Tukey HSD (*Table 20*), shows this significant difference, as well as statistically significant differences between English secondaries and Scottish primaries:

Table 20: Dysfunction Index - Tukey HSD

(I)	(J)	MEAN DIFFERENCE (I-J)	STANDARD ERROR	Sig.	95% CONFIDENCE INERVAL FOR MEAN	
					Lower	Upper
English Secondary	2	.5435*	.1304	.000	.2085	.8785
	3	-.0821	.1500	.947	-.4674	.3032
	4	-.4765*	.1462	.006	-.8520	-.1011
English Primary	1	-.5435*	.1304	.000	-.8785	-.2085
	3	-.6256*	.1584	.000	-1.0325	-.2188
	4	-1.0201*	.1548	.000	-1.4176	-.6225
Scottish Secondary	1	.0821	.1500	.947	-.3032	.4674
	2	.6256*	.1584	.000	.2188	1.0325
	4	-.3944	.1716	.098	-.8352	.0463
Scottish Primary	1	.4765*	.1462	.006	.1011	.8520
	2	1.0201*	.1548	.000	.6225	1.4176
	3	.3944	.1716	.098	-.0463	.8352

* Significant at the .05 level.
1 = English Secondary, 2 = English Primary, 3 = Scottish Secondary, 4 = Scottish Primary

And not too surprisingly an analysis of homogeneity, *Table 21*, also shows these differences as well as the similarities between the secondary schools from the two countries:

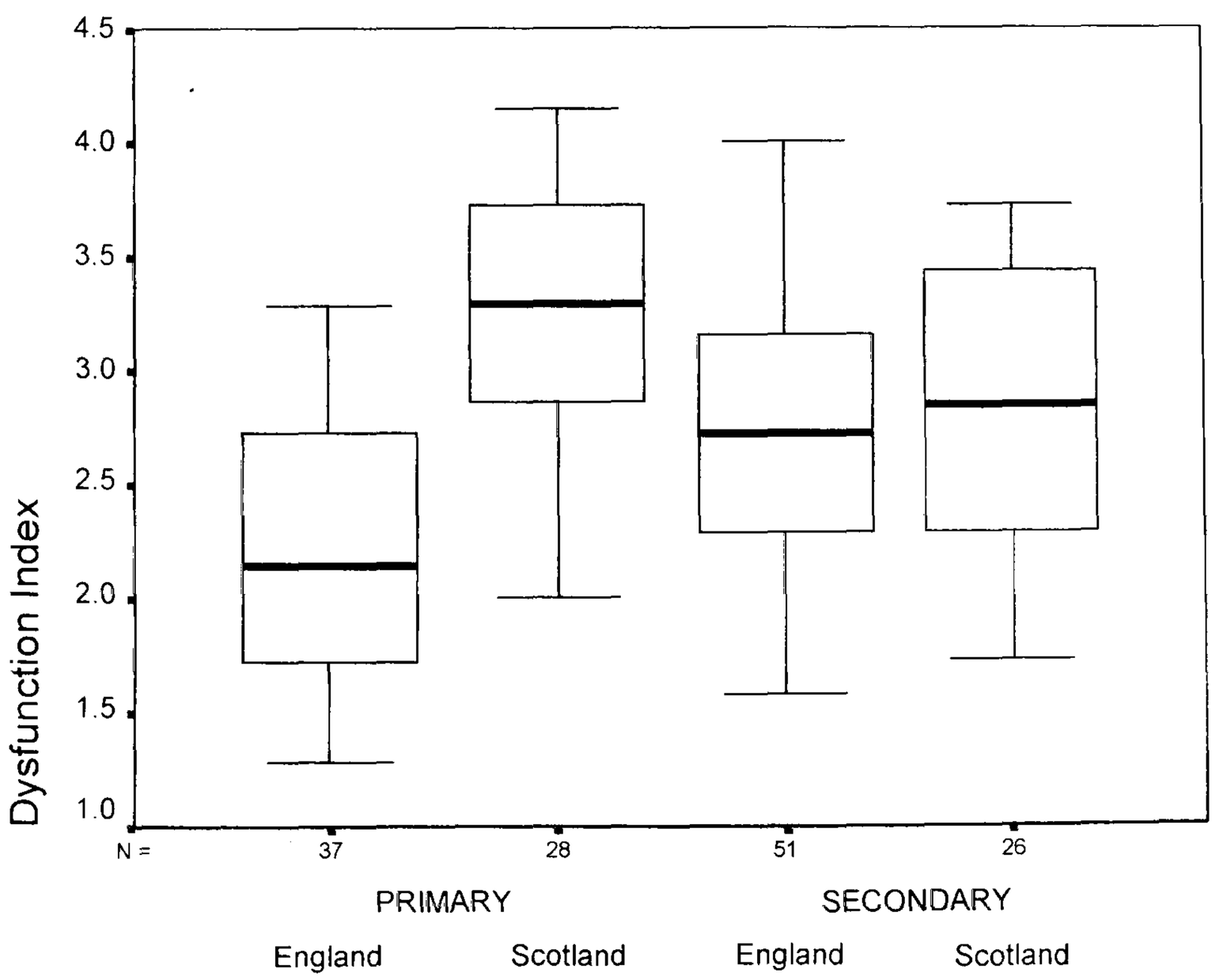
Table 21: Dysfunction Index - Homogenous Subsets

	N	SUBSET FOR ALPHA = .05		
		1	2	3
English Primary	39	2.1757		
English Secondary	53		2.7192	
Scottish Secondary	26		2.8013	
Scottish Primary	28			3.1958
Sig		1.000	.950	1.000

Means for groups in homogeneous subsets are displayed
a Uses Harmonic Mean Sample Size = 32.909

Furthermore, these differences are very evident in box plots (*Figure 10*) which based on rankings of the dysfunction indices;

Figure 10: Dysfunction Index - Country / Phase



Notes: The lower the index number the greater the dysfunction

With respect to the analyses carried out for H3 (secondary – primary) it is clear from the above that this was not very satisfactory. The fact that little significant difference was found in the responses to the KPI – Dysfunction questions, was very much due to the responses from the English and Scottish primary schools 'cancelling' each other out.

In view of this shortcoming it was decided to create and test two more hypotheses which would allow these differences to be tested. They are primarily derived from H3 and H4:

SH1 – English and Scottish primary schools have similar views on the use of performance data for school management, but the English schools find KPIs more dysfunctional. This was expected to be true because schools in both countries make extensive use of data from their own sources and nationally based tests. However, it was expected to find that the KPIs in England have more of a dysfunctional effect, because of their greater prominence and the publication of performance and league tables.

SH2 - English and Scottish secondary schools have similar views on the use of performance data for school management, and on the dysfunctional effects of KPIs. This was expected to be true because secondary schools in both countries widely use performance data, and the KPIs are published in both countries.

SH1 English and Scottish primary schools

This hypothesis looks at the differences between the English and Scottish primary schools (Heads and Teachers). It was expected that the respondents (*Table 22*) would have similar views on the use of performance data in the organisation and management of schools.

Table 22: KPIs – Management: English and Scottish primary schools

QUESTION	MEAN		STANDARD DEVIATION		SIGNIFICANCE	
	England	Scotland	England	Scotland	t-Test p value	Effect size
Numeric Data	2.35	2.61	1.11	1.20	.400	0.23
Targets Useful	2.89	2.52	1.10	1.12	.188	-0.33
Parents	4.08	4.07	0.91	1.04	.970	-0.01
Pressure Targets	2.08	1.67	0.89	0.78	.059	-0.49
Participate Targets	1.41	1.41	0.80	1.34	.992	0.00
Judge Performance	2.86	2.88	0.92	1.15	.926	0.02

Professional Judgements	3.78	3.26	0.98	1.13	.051	-0.50

Mean: A score of 1.00 would represent all respondents strongly AGREEING with the question or the first term in the semantic differential
A score of 5.00 would represent all respondents strongly DISAGREEING with the question or the second term of the S.D.
England N = 39 Scotland N = 28 (Heads and Teachers)

Overall the responses from these two groups are very similar. However, probably the most interesting finding is that the differences found in H4 between English and Scottish respondents in terms of pressure to meet targets appear to remain. The p value is no longer statistically significant (at the .05 convention) [149] although the effect size has increased from –0.39 to –0.49. It is worth noting that this level of pressure is similar across both phases in Scotland, with mean values of 1.77 and 1.67 (p=.537).

As discussed for H4 this would seem to be an important finding and one that can not be easily explained from either the data or additional comments; and would therefore benefit from further investigation. As a consequence of this finding it may well be that primary league tables in England do not put any more pressure on the schools than other, less obvious and less public methods used in Scotland. And it follows that public 'name and shame' policies, which are inextricably linked to league tables, would seem to be of little real benefit.

The finding that there was a difference between the reference parents made to KPIs in Scottish and English primaries was also surprising. Given the greater prominence KPIs and league tables have in England, it had been assumed that they would encourage greater discussion between schools and parents. Again this questions whether league tables are in fact doing what they are supposed to; namely encourage parents to put pressure on schools to improve performance.

Nevertheless, these findings do largely support the first part of the hypothesis, in that it would appear that English and Scottish primary school have similar needs for performance data, and use this data in similar ways. The second part, relating to potential dysfunctional effects is considered below, and it was expected that English primary schools would find KPI systems more dysfunctional.

Table 23: KPIs – Dysfunction: English and Scottish primary schools

QUESTION	MEAN		STANDARD DEVIATION		SIGNIFICANCE	
	England	Scotland	England	Scotland	t-Test p value	Effect size
Harmful Overall	2.49	3.70	1.04	1.35	.000	1.03

[149] It would be reasonable to surmise that with a few more cases the .05 (target!) would be met

Conflicting Aims	2.71	3.78	1.00	0.97	.000	1.08
Concentrate Targets	2.56	3.37	1.21	1.01	.001	0.72
Narrow Curriculum	1.25	2.19	0.55	0.92	.000	1.33
Borderline Children	2.64	3.67	1.27	1.18	.006	0.84
Blame Culture	1.51	2.50	0.73	0.95	.000	1.20
Undermining	2.08	3.07	1.26	1.04	.002	0.85
INDEX	2.17	3.19	0.57	0.65	.000	1.68

Mean: A score of 1.00 would represent all respondents strongly AGREEING with the question or the first term in the semantic differentials
A score of 5.00 would represent all respondents strongly DISAGREEING with the question or the second term of the S.D.
England N = 39 Scotland N = 28 (Heads and Teachers)

The differences between these responses are very striking. In all cases the English primary schools view their KPIs as substantially more dysfunctional than their Scottish counterparts. Given the significance of the findings the individual items are discussed in more detail below.

Harmful Overall: Although *Survey 6* showed that all of the respondents combined found KPIs mainly beneficial, this comparison paints a very different picture. Many English primary schools felt that KPIs were harmful, and this is particularly significant when considering that they agreed, to a similar degree, that schools need performance data, and that target setting is useful.

Conflicting Aims: This considers the effect of target setting at the whole school level, by aiming to assess the degree of conflict between the targets and the school's aims, and as previously pointed out (*Survey 15* and H5) in general most of the respondents did not feel that there was much conflict. However this finding again shows a very different picture for English primary schools. Potentially it questions the value of strategic or school development planning, which starts with the aims, and leads on to objectives and targets (see chapter 4), if as might be the case, school development planning starts with the targets. A further question (*Survey 16*) suggested that English primary respondents were more likely to find that if there was conflict, then the targets would override the aims.

Concentrate Targets: This question considers KPIs at the school operational level.
Most people would agree that schools have many objectives in addition to those prescribed by the targets. The last chapter pointed to the apparent split in responses (*Survey 22*), and this may well be partly explained by the differences reported here. These results would suggest that English primary schools are more likely to concentrate on their targets at the expense of, what the respondents consider to be, other important objectives. This raises important questions about the long term

value of education offered by English primary schools, if for example learning or social skills objectives are not met. Furthermore, this may have a negative effect on the next stage of education.

Narrow Curriculum: This looks more specifically at the subjects that are actually taught as part of the curriculum in the school. There was very strong support (*Survey 8*) for the notion that KPI systems have a narrowing effect, however this finding is even more extreme. It suggests that particularly in English primary schools, subjects which are not assessed by the KPIs, such as music, PE and technology, are given less prominence in the curriculum. As has been previously pointed out, this effect of KPIs had been widely predicted and observed. There are many implications arising from this, including longer term health issues, and the technological ability of future citizens.

Borderline Children: This considers the issue at the classroom level, and the attention which particular groups of children receive. *Survey 60* showed that the majority of schools were targeting 'borderline' children, however this analysis shows the significant difference between primaries in the two countries. As previously pointed out there may of course be perfectly valid reasons for certain groups receiving extra help, but if the main purpose behind this targeting is to help schools (or the government) meet their targets, this then would seem to be dysfunctional.

Blame Culture: This question considers the degree to which KPI systems encourage a culture of blame. *Survey 29* showed that this was an issue for many of the schools, however for many English primary schools it does appear to be much more of a problem. 62% of English respondents 'scored' 1, the extreme end of the semantic differential, as opposed to 13% from Scottish primaries. An ANOVA post Hoc Tukey HSD analysis of homogeneity illustrates the significant differences between English primary schools and the others:

Table 24: Analysis of homogeneity for Survey 29 – Blame culture

	N	SUBSET FOR ALPHA = .05	
		1	2
English Primary	39	1.5128	
Scottish Secondary	24		2.1667
English Secondary	53		2.2830
Scottish Primary	28		2.4955
Sig		1.000	.485

An increase in blame culture may well have many serious and long terms dysfunctional effects on any organisation, in particular primary schools. An effective school by definition needs a very high degree of trust and co-operation between staff, and in increase in 'blame' is bound to have a negative effect. This finding very much contradicts the government's view that there is now less of a blame culture at all levels, because of the educational reforms (Barber 1999).

Undermining: This final question considers whether respondents themselves find that KPI systems are undermining or supportive of their work. The Scottish primary responses were relatively neutral, however for their English counterparts this would appear to be quite a significant problem. Given their support for using performance information (Numeric data) in general, it would seem that the KPI systems may not necessarily be doing what was intended, namely supporting heads and teachers.

Overall this analysis very strongly support the hypothesis that English primary schools find their KPI systems more dysfunctional than Scottish primaries. The analyses in the last chapter showed that many of the issues were a cause for concern for all the schools in the survey, however what is evident here, and must be of even more concern, is the far greater dysfunctional effects on primary schools in England.

SH2 English and Scottish Secondary schools

This hypothesis looks at the differences between the English and Scottish secondary schools (Heads and Teachers). It was expected that the respondents (*Table 25*) would have similar views on the use of performance data in the organisation and management of schools.

Table 25: KPIs – Management: English and Scottish Secondary schools

QUESTION	MEAN		STANDARD DEVIATION		SIGNIFICANCE	
	England	Scotland	England	Scotland	t-Test p value	Effect size
Numeric Data	1.45	1.84	0.50	0.90	.053	0.58
Targets Useful	2.32	2.54	1.19	1.07	.432	0.19
Parents	3.71	3.92	1.13	1.26	.455	0.18
Pressure Targets	1.96	1.77	0.79	0.71	.299	-0.25
Participate Targets	1.26	1.31	0.49	0.55	.721	0.10
Judge Performance	3.10	2.96	1.12	0.97	.606	-0.13

Professional Judgements	3.60	3.92	0.86	0.74	.110	0.40

Mean: A score of 1.00 would represent all respondents strongly AGREEING with the question or the first term in the semantic differential
A score of 5.00 would represent all respondents strongly DISAGREEING with the question or the second term of the S.D
England N = 53 Scotland N = 26 (Heads and Teachers)

At the secondary level, (as with the primary schools), it appears that there is little difference between Scotland and England, in terms of how they use KPI systems for management. The difference in the use of Numeric data was on the borderline of statistical difference in terms of p value, there was also quite a sizeable difference in ES. This is consistent with the findings from H4, and as suggested before this may in part be due to the greater prominence of KPIs in England and the use of 3rd party indicator systems; all of which might encourage English schools to use data more.

Taken together these findings do largely support the first part of the hypothesis, in that it would appear that English and Scottish secondary schools have similar needs for performance data, and use this data in similar ways. The second part, relating to potential dysfunctional effects is considered below, it was expected that there would be little difference between English and Scottish secondary schools.

Table 26: KPIs – Dysfunction: English and Scottish secondary schools

QUESTION	MEAN		STANDARD DEVIATION		SIGNIFICANCE	
	England	Scotland	England	Scotland	t-Test p value	Effect size
Harmful Overall	3.25	3.35	1.30	1.26	.756	0.08
Conflicting Aims	3.72	3.46	1.03	1.07	.308	-0.25
Concentrate Targets	3.26	2.84	1.09	1.29	.137	-0.36
Narrow Curriculum	2.06	2.46	1.05	1.27	.137	0.35
Borderline Children	2.15	2.58	1.06	1.25	.122	0.38
Blame Culture	2.28	2.17	0.97	0.92	.621	-0.12
Undermining	2.34	2.72	1.25	1.24	.214	0.31
INDEX	2.72	2.80	.65	.60	.593	0.13

Mean: A score of 1.00 would represent all respondents strongly AGREEING with the question or the first term in the semantic differentials
A score of 5.00 would represent all respondents strongly DISAGREEING with the question or the second term of the S.D.
England N = 53 Scotland N = 26 (Heads and Teachers)

This analysis shows very little difference between the two groups, and with no statistical difference between any of the items. This very much supports the hypothesis that there is little difference between the views of English and Scottish secondary schools in terms of both the use of PIs and

their potential dysfunctional effects. From this it is suggested that the National and cultural differences do not significantly affect the respondent's views.

Conclusion

In this chapter a number of comparative analyses were carried out between different groups of respondents, with the aim of identifying differences and similarities in the responses, which might in turn be able to help understand some of the issues surrounding PIs. Three hypotheses were tested; H3 (English – Scottish schools), H4 (Primary – Secondary), and (H5) Heads – Teachers. However, it became apparent that the findings from the first two would not satisfactorily explore some of the issues emerging in the data, therefore two supplementary hypotheses were formed; SH1 (English primary – Scottish primary), and SH2 (English secondary - Scottish secondary). The analyses were performed with two groups of questions from the main questionnaire. The first looked at how PIs are used in the management of schools, and the second how they may have dysfunctional effects. In addition an index of dysfunction was formed from the latter group.

In terms of the first set of questions there was quite strong evidence that most of the respondents from the different groups, used and viewed PIs in similar ways. For example, all of the groups were in favour of the use of targets (*Survey 27*). However, one important difference and particularly interesting finding was that Scottish schools (primary and secondary) appeared to be under greater pressure to meet their targets than their English counterparts. There were no obvious explanations for this, either in the data or from the various additional comments. It was suggested here that the most likely explanation was the different relationships schools had with their local authorities, north and south of the boarder.

With respect to the second group of questions (KPIs – Dysfunction) the most dramatic findings were the differences between the English and Scottish primary schools (SH1). For all of the items in this group the responses from the English primaries suggested that their KPIs were significantly more dysfunctional. Issues of national and cultural differences which may have had some effect on the responses, were not felt to be responsible for the very significant and large differences. This is mainly because many of these national and cultural issues would have had similar effects on the secondary schools in both countries, and for these very little difference in the responses were found.

The key organisation difference which was felt to be largely responsible for these findings was the publication of performance data (league tables), which occurs in England but not in Scotland. These it was suggested would encourage schools to engage in dysfunctional behaviour. It is

however worth noting that the Scottish primaries did feel under greater pressure to meet their targets, and this suggests that publishing data does not necessarily increase pressure on schools to meet targets.

The final hypothesis looked at the differences between heads and teachers. The findings suggested a very high degree of agreement between these two groups, with virtually no significant differences in the responses. On the face of it this must be good - with such a high apparent degree of harmony between 'managers' and 'managed': however some degree of caution should be applied. The greatest threat to validity was in the selection of the teacher respondents. Because of the method of randomisation there was a possibility of some degree of self selection or 'head selection', which might have led to an sample of teachers, who were unrepresentative of all teachers, and being 'too' much in agreement with their heads. This is however not to suggest that these findings are necessarily wrong, but rather to propose that more corroborative evidence in needed.

One analysis worthy of mention in this conclusion, which was not reported in the main part of this chapter, was the differences between English primary and secondary schools. Again little difference was found in terms of the usage of PIs, however significant differences were found for the dysfunctional effects (primaries greater). The dysfunction index gave means of 2.75 and 2.17, ($p<.001$, ES=0.82). In both cases data is published in the form of high profile league tables, furthermore many of the organisational structures, which might have had an influence on the responses, are common to both (eg. Ofsted, LEAs and LMS / pseudo market principles). Two possible explanations are tentatively proposed; firstly league tables and some aspects of LMS are a more recent phenomena for primary schools and these are simply bedding in, and in due course the associated KPIs will have a less dysfunctional influence. Secondly, primary schools are by definition smaller and this will make their results appear far more inconsistent, due to natural fluctuations in the cohort. Furthermore, primary heads have to take a more hands on approach to the school and the teaching, and this may make them more aware of dysfunctional influences.

This chapter has shown up a number of very important differences (and similarities), which raise serious questions about the effect of KPIs and the associated league tables; and in particular their effect on English primary schools. The next chapter looks at some relationships (correlations) in the data, which may further help identify some of the effects of KPIs.

Chapter 13

Follow up survey and relationships between responses

Faith is the great cop-out, the excuse to avoid the need to think and to evaluate evidence. Faith is belief in spite of, or even perhaps because of, the lack of evidence. (Dawkins)

Introduction

The are two parts to this chapter, the first looks at how the views of the respondents may have changed over time, and the second looks for relationships between the responses to some of the questions. At the time of the main survey a number of quite significant changes in the organisation of the education systems were happening or were under discussion, in both England and Scotland, and these may have affected some of the responses. In addition responses to some questions may have been affected by other factors, for example how well a school performs may have a bearing on how dysfunctional it finds its KPIs.

Follow up survey

This was based on survey carried out about a year after the main survey, with a number of respondents who had indicated a willingness to be contacted (about 40% offered). This included heads, teachers and governors. It was decided, however, to limit the follow-up survey to heads, primarily because they were the easiest to contact, but also to try and compare 'like with like'. The smallest group of willing respondents was from Scottish secondary schools (also least number surveyed), with eleven potential participants. This group was contacted first, and from this eight positive contacts were made, (two had left for other posts and one had retired). For the other schools the potential participants were listed in random order, and this was followed until eight contacts were made for each type of school. The survey itself took the form of a semi-structured telephone interview, (*Appendix 7*), which aimed to maintain a degree of consistency and objectivity between responses. For most of the respondents this was followed by a general discussion about the issues considered by the thesis.

The fourteen questions from the two themes (*KPIs Management* and *KPIs Dysfunction*) were re-asked. The questions were re-ordered with the first nine (survey 72, 79, 2, 13, 19, 38, 22, 19, 15) being in the same format as the original questionnaire, (five point likert scale). The other five semantic differentials (survey 6, 29, 8, 27, 9) broadly followed the format in the questionnaire, although respondents were asked to score their responses on a 1-5 scale.

The aim of this follow up survey was primarily to assess how the respondents perceptions had changed over the year, and not to draw any specific conclusions or make generalisations from the particular answers. The results are presented below by comparing the respondents' questionnaire data with their follow up telephone data. Clearly, the sample (32 in all) are unlikely to be very representative, in part because they are essentially self-selecting. That said, it is however worth

noting that there was no correlation (Corr. 0.009, p > .05, n=162) between those agreeing, or not agreeing to take part in, the follow-up and their Dysfunction index scores. Put another way the follow-up sample did not appear to be significantly different to the larger sample.

KPIs and the management processes

Table 27: KPIs – Management: 1999 and 2000

QUESTION	MEAN		STANDARD DEVIATION		SIGNIFICANCE		CORRELATION 1999 - 2000	
	1999	2000	1999	2000	t-Test	ES	Corr.	Sig.
Numeric Data	1.94	1.84	1.05	1.11	.729	-0.09	.658	.000
Targets Useful	2.25	2.53	0.98	1.27	.326	0.25	.691	.000
Parents	4.19	3.09	1.05	1.25	.000	-0.97	.552	.001
Pressure Targets	1.75	1.78	0.88	0.97	.893	0.03	.423	.016
Participate Targets	1.22	1.13	0.55	0.34	.415	-0.20	.195	.284
Judge Performance	3.28	3.28	1.02	0.99	1.00	0.00	.428	.014
Professional Judgements	3.63	3.25	1.16	1.11	.190	-0.40	.227	.125

Mean: A score of 1.00 would represent all respondents strongly AGREEING with the question or the first term in the semantic differential
A score of 5.00 would represent all respondents strongly DISAGREEING with the question or the second term of the S.D.
1999 (Questionnaire) N = 32 2000 (Telephone) N = 32

On the whole there was close agreement between the two years, (and survey methods). The difference in terms of parents making reference to KPIs is interesting. A further analysis of this shows that it is the English secondary schools which had changed most significantly (means 4.13 and 2.59, sig .004, ES 1.86, n=8). Clearly the numbers are very small and no firm conclusions should be drawn, however a possible explanation might be the timing of the two surveys. The first was in June 1999 and the second in the September 2000. The archives of The Daily Telegraph (www.telegraph.co.uk) and Schools Net (www.schoolsnet.com) were consulted and the most significant articles in September 2000 not present in June 1999 related to the 'A' level league tables, (also teacher shortages) which might have prompted parents to discuss KPIs.

Potential dysfunctional effects of KPIs

Table 28: KPIs – Dysfunction: 1999 and 2000

QUESTION	MEAN		STANDARD DEVIATION		SIGNIFICANCE		CORRELATION 1999 - 2000	
	1999	2000	1999	2000	t-Test	ES	Corr.	Sig.

Harmful Overall	3.25	3.19	1.32	1.33	.851	-0.05	.652	.000
Conflicting Aims	3.44	3.22	1.01	1.16	.424	-0.20	.493	.004
Concentrate Targets	3.06	3.38	1.19	1.13	.285	0.27	.510	.003
Narrow Curriculum	1.94	2.31	0.95	0.97	.122	0.39	.445	.011
Borderline Children	3.09	2.56	1.23	1.39	.110	-0.41	.308	.086
Blame Culture	2.31	3.09	0.97	0.86	.001	0.84	.705	.000
Undermining	2.31	3.28	1.15	0.85	.000	0.95	.402	.023
INDEX	2.77	3.00	0.64	0.61	.145	0.37	.700	.000

Mean: A score of 1.00 would represent all respondents strongly AGREEING with the question or the first term in the semantic differential
A score of 5.00 would represent all respondents strongly DISAGREEING with the question or the second term of the S.D.
1999 (Questionnaire) N = 32 2000 (Telephone) N = 32

Two issues stand out; firstly the reduction in 'blame culture' and KPIs being more supportive (ie less undermining). The main 'contributor' to the first was the English primary schools (means 2.00 and 3.25, p<.001, ES. 1.26, n=8). On the face of it, and if true, a very pleasing change, although there were few clues as to why this should have occurred[150]. The second significant change was mainly due to the Scottish secondary schools (means 2.13 and 3.38, p<.001, ES. 1.30, n=8), and again no obvious clue as to the cause, although it is perhaps interesting to note that in Scotland at that time, there were widely reported problems with issuing of results and grades for the highers.

Table 29 look more specifically at how the responses changed for the different types of schools, and this appears to confirm a small (and strictly speaking not statistically significant) trend towards KPIs being less dysfunctional.

Table 29: Dysfunction Index – Changes from 1999 to 2000

QUESTION	MEAN		STANDARD DEVIATION		SIGNIFICANCE	
	1999	2000	1999	2000	t-Test p value	Effect size
English Secondary	2.52	2.84	0.73	0.60	.270	.47
English Primary	3.11	3.30	0.59	0.53	.523	.34
Scottish Secondary	2.86	3.02	0.68	0.63	.639	.24
Scottish Primary	2.61	2.86	0.51	0.68	.459	.43

Mean: Dysfunction Index, the lower the score the greater the reported dysfunction. N = 8 for all cases

[150] It is tentatively suggested later that the survey method may had some influence, as respondents may feel inhibited to say on the phone that there is a blame culture in 'their' school

Conclusion

Overall, it would appear that the findings have remained broadly stable. Given the small (and self-selecting sample) it would be wrong to read too much in to the specific differences. With respect to the overall 'improvement' in dysfunctional scores, this may simply be explained by the differences in the survey methods, in that respondents may be more inhibited in giving a negative impression over the phone. On the other hand this may also be an indication that the many 'new' initiatives have bedded down and schools are now better able to cope with the potentially dysfunctional effects of KPIs.

Relationships between the results

This section looks at some of the relationships (correlations) between the answers to the survey questions, which may help explain some of the findings. The analysis was carried out with SPSS (Pearson co-efficient)[151]. Scatter graphs were plotted on screen to get an idea of the distributions and check for curvilinear relationships. The correlation co-efficients are reported and in terms of describing what they signify, Fink (1995) offers the following view[152]:

0.00 - 0.25	=	Little or no relationship
0.26 - 5.00	=	Fair degree
0.51 - 0.75	=	Moderate to good
0.76 - 1.00	=	Very good relationships

(p. 36)

She is however at pains to point out that this is very much dependant on the nature of the survey questions and particular circumstances. In the case of the questions considered in this research, they are not based on objective or actual physical measurements, but rather individual perceptions, and this will reduce the likelihood of achieving very high levels of correlation. Furthermore, the individuals concerned are not a particularly homogenous group, but a wide collection of individual heads and teachers with many different ideas and feelings towards their schools, and the education system more generally. In addition there may well be differences in terms of interpreting and understanding the particular questions, therefore fairly low correlation coefficients may be significant.

The co-efficient values are reported along with the statistical significance (p value) and the number of cases. Where appropriate, scales have been reversed (eg. other end of SD) to clarify the relationship and to avoid double negative statements. Three key themes have been identified as

---

[151] Spearman ranks were also calculated to check for any significant differences which might have been missed
[152] Described as a "Conservative rule of thumb".

being of particular relevance to the issues considered in this thesis: firstly the effect of respondents perception of their relative performance: secondly, their general view towards performance data, and thirdly, how useful they find targets and target setting. In addition, all of the questions were analysed to see if any other interesting or potentially significant relationships 'popped up'.

Relative Performance

This would seem to be a relevant and interesting variable: in that how views towards performance indicators might be affected by a school's relative performance. The variable used for comparison was the introductory question: *How would you assess your school's position to others in the LEA Local area.* The overall responses were discussed in chapter 10 and the distributions are shown in *Figure 8.* As pointed out this question was concerned with the respondents' perceptions of their schools performance, rather then a precise measure of performance, either at the particular time or over a period of years.

Relative performance – How fair: (*Figure 8* and *Figure 9*) There was quite a strong and significant positive correlation between the two questions: (Corr. 0.623, p =.000, n=143). Whilst it is hardly surprising that 'good' schools find the systems (KPIs) by which they are judged fair, the serious implication for the education system overall, is that schools which do not perform very well, possibly for reasons beyond their control, may well have less trust and confidence in the measures by which they are judged. Leading on from this, it was found that 'high' performing schools were less likely to indicate that the statutory targets should consider more aspects of their school.

Relative performance – More aspects (*Figure 8* and *Survey 20*). (Corr 0.191 p=.017 n=159). Arguments have already been put forward that the KPIs should be composed of a more balanced set of indicators, however this suggests that better performing schools would be less in favour of such an approach.

Relative position - Dysfunction index: (*Figure 8* and *DI*) A logical and tempting explanation for many of the results would be that schools which judged their performance to be low in terms of KPIs would feel that those systems were dysfunctional, and those that performed well would feel the opposite. However, this was not found to be the case, with some correlation (Corr 0.232, p=.004, n=155) suggesting that high performing schools reported more dysfunctional behaviour: a rather surprising result. The analysis was checked and confirmed in a number of ways. Firstly, for the different types of schools, *Table 30* shows some degree of positive correlation for all except Scottish secondary. Interestingly, this relationship is strongest for Scottish primaries (which had the least dysfunction), suggesting that 'low' performing schools in this sector are very 'un-

dysfunctional'. An ANOVA analysis confirmed there was no significant difference between the reported relative position means of each type of school.

Table 30: Correlation - Relative performance and Dysfunction Index

	CORR.	SIG.	N=
English Secondary	.294	.021	61
English Primary	.197	.194	45
Scottish Secondary	.080	.703	25
Scottish Primary	.400	.053	24
ALL SCHOOLS	.232	.004	155

Secondly, the correlation between the various components of the dysfunction index and the relative performance were calculated (*Table 31*). In all of the cases the positive relationship held:

Table 31: Correlation - Relative KPI performance and Dysfunction Index components

QUESTION	CORR.	SIG.	N=
Harmful Overall	.122	.130	156
Conflicting Aims	.086	.287	157
Concentrate Targets	.166	.039	155
Narrow Curriculum	.078	.332	156
Borderline Children	.137	.090	153
Blame Culture	.184	.022	155
Undermining	.199	.013	156
DYSFUNCTION INDEX	.232	.004	155

All responses from the survey

The conclusion from this analysis, perhaps rather disturbingly, is that 'high' performing schools are more likely to find their KPIs systems, by which they have been judged 'high' performing, more dysfunctional. And perhaps more positively, it would seem that 'low' performing schools are less affected by dysfunctional influences.

Relative position – Parents discuss (*Figure 8* and *Survey 2*) Some positive and significant correlation was found between a schools performance and how much reference parents made to the KPIs (Corr 0.277, p=.005, n=154). A simple and quite likely explanation is that higher performing

schools have 'higher performing' parents who are more confident and likely to discuss the school's performance. And of course in high performing schools these discussions will tend to be positive. which may in itself be a reason to discuss them.

However, for low performing schools it would seem that parents are less likely to discuss school performance, which further questions the principle of parental accountability (eg. parents charter). and parents 'forcing' schools to improve. Related to this there was little evidence that schools that did not perform so well were under greater pressure to meet targets, than their high performing counterparts; little (if any) correlation was found between a schools relative position and pressure (*Figure 8* and *Survey 13*) (Corr 0.151, p=.060, n=156).

Attitude towards performance data

Survey 72 showed the widespread support for good objective performance data, and this appeared not to be dependant on either a schools performance (Corr .060, p=.458, n=156) or degree of dysfunction (Corr -.016, p=.841, n=155). Furthermore, the length of time in post also did not appear to be related to the need for performance data (Corr -.118, p=247, n=158). It had been expected that newer heads would be more in favour of performance information and target setting than the 'old school'. Indeed, the length of time in post did not significantly correlate with any of the other findings, including perceived dysfunction (Corr .066, p=.519, n=155).

Targets useful

Survey 27 showed a good deal of support for targets and target setting, and not surprisingly this correlated positively with the view that this would help raise standards (Corr .562, p=000, n=159). Those schools which valued target setting more, were less likely to feel that KPIs encouraged central control[153] (Corr .219 p=.006, n=159). However, there was no evidence in terms of statistical correlation that 'high' performing schools were more or less enthusiastic about targets and target setting (Corr .089 p=.269, n=158). Although, as would be expected there was some correlation between attitudes towards targets and the need for good performance data (Corr .190 p=.017, n=159).

Conclusion

The findings from this chapter do suggest that the results from the first questionnaire survey are valid in terms of the time scale, in that no significant differences were found when the same questions were asked to the same respondents a year later; and this is in spite of quite a number of significant changes during this period, such as the development of performance management

---

[153] Survey 12

systems in England and the McCrone committee report in Scotland. The reported general 'improvement' or reduction in dysfunctional behaviour, is however, mainly attributed to the different survey methods, although it would be interesting to know what the responses would have been to another anonymous postal questionnaire. Some degree of 'bedding down' and increased tolerance towards the negative impact of KPIs was alluded to by several of the respondents.

In terms of the relationships between the data, as would be expected schools which performed well, felt more positive towards the indicators by which they are judged. The finding that a school's relative performance did not affect its view of KPIs was felt to be positive. However, the finding that 'better' performing schools were more likely to report / engage in, dysfunctional behaviour was surprising, and therefore this did not support the hypothesis (H6). This is quite an alarming conclusion, but in retrospect perhaps not so surprising, as it suggests that for schools to be 'successful' they will have to engage in more dysfunctional behaviour.

Chapter 14

Conclusions

League tables are too crude a measure for schools such as Gordonston (HMC)

Introduction

This concluding chapter firstly provides a summary of the previous chapters which make up the thesis; secondly, it looks at five key themes, *targets, accountability, complexity, the right measures and broader organisational context*; which run throughout the thesis; thirdly, a brief look at how KPI systems may develop in the future; and finally, a critique of the approach taken by the thesis.

Summary

Chapter 1 discussed the increasing use of performance indicators in schools, which has occurred over the last twenty or so years, and in particular the issue of 'high stakes' Key Performance Indicators. It was however suggested, that their introduction had not been accompanied with appropriate consideration about their effects on the organisation and management of schools; and more crucially their potential dysfunctional consequences.

The broad plan for the thesis was discussed. This identified four main elements, which were to be used to explore and where appropriate quantify the effects of KPIs on English and Scottish schools. The first element looked at the place of PIs in the broader organisational context, and made particular reference to generic management theory. The second aimed to consider some aspects of management accounting theory which is relevant to high stakes PIs. The experiences of other public sector organisations with KPIs was considered in the third section. And in the final part a research program aimed to identify and assess a number of the issues surrounding KPIs and schools.

Chapter 2 started by discussing what PIs are and how they are used in the accountability and management processes. The inherent tensions between these two different uses of PIs was discussed as was the complex nature of the education process itself. This complexity makes predicting the impact of PI systems very difficult, and is further complicated by the 'high stakes' nature of the indicators, and the broader political context within which schools operate.

It was suggested that whilst the government strongly believe in the value of PIs to help manage the education system and to raise standards generally, there are however, inconsistencies in their approach. On the one hand it is claimed that PIs should be used in a broad and balance way, and to develop a 'learning organisation' culture; on the other they appear to be used to impose targets, which often seem arbitrary and disjointed. This confusion and related contradictions are evident in many of the issues considered by this thesis.

The important issue of making fair and reasonable comparisons between schools was explored, and from this perspective, the inadequacy of league tables based on raw data was discussed. Ways of improving the tables were considered, such as value added and multi-level modelling. Finally, the issue of who manages 'what', in the overall system was considered. It was argued that although schools with the advent of LMS are supposed to be autonomous, in reality they are still tightly organised and controlled by central government, both in England and to a lesser extent in Scotland (EAs still have an important role).

Much use has been made throughout the thesis of theory and experiences from other sectors, and the case for this approach was made in chapter 3. In spite of what might appear to be the mistrust of business theory by some in education, a number of academics have used such theory to good effect. It was concluded that in essence businesses, just like schools, are human organisations and have much in common in terms of motivation, control and performance assessment. Indeed, evidence of a convergence in the theories was discussed, and the increasing privatisation of educational services will further erode any differences. However, it should be noted that no judgement was made in favour or against the privatisation trend, and although these issues are important, and possibly in an indirect way of significance to this thesis, such considerations are beyond its scope.

Chapter 4 identified three different schools of organisational theory, which it was felt were most relevant to the issues considered here. The first, *Rational and Scientific* looked at the principles of Bureaucracy, Scientific management and Planning. It was evident from the accompanying discussion that much of the educational management and organisation follow these rational / scientific principles. The education system is still very much a hieratical organisation, with for example, great emphasis being placed on the close association between a 'strong head' and an effective school. Planning is of great importance to schools and L(EA)s and this very much follows bureaucratic principles. Many of the actual teaching processes are essentially 'scientific methods', being based on the claimed 'one best way' principle; for example, the national literacy / numeracy strategies. Although it should be noted, that proper scientific methods, (eg. RCTs) are not generally used to identify the actual methods.

The second school, the *Behaviourist*, considered the work of Elton Mayo and the Hawthorne studies, as well as Douglas McGregor's theory X and Y. Mayo's work is very important to this thesis as it was the first major study on the softer issues of management, in particular the behavioural consequences of particular management and organisational interventions. McGregor's work is interesting in that theory X provides a largely pessimistic view of workers and presumes that they dislike work and need to be tightly organised, with for example performance related pay

and performance management systems. Conversely, theory Y takes an optimistic view and believes that workers are self motivated and willing to perform well and to the best of their abilities. Theory Y largely represents the government's espoused views, however for many X is nearer the reality.

The final school, *Hybrid*, is largely built up with elements from both the behaviourist and the rational scientific schools, and represents much of 'modern' management theory. Principles such as management by objectives and Peter's and Waterman's ideas on excellence were used to introduce three approaches; Total Quality Management, Learning Organisations, and Business Process Reengineering. Elements from these are quite common in the educational management literature. Whilst TQM as a philosophy has much to offer education, in particular with respect to the use of PI systems, its effectiveness depends very much on how it is implemented. For example, what was referred to as 'hard TQM' was discussed, and in effect this is little more than a crude implementation of scientific management.

The learning organisation approach has many appeals, not least the name, and again if used and applied properly, would seem to be of benefit to any organisation, and most certainly schools. However, the key as the name suggests, is that learning takes place throughout the organisation. Whilst many schools are in themselves able to learn, the problems would seem to be one of learning going up the system, leading ultimately to the whole national education system becoming a learning organisation.

The third approach BPR has a number of attractions, in particular to organisations seeking dramatic or 'step' improvements in their performance (especially that measured by KPIs). In essence this would appear to be the government's preferred approach to 'improving' the public sector, including education. However, it was argued that the long term effectiveness of such approaches are of limited value to complex organisations such as schools; furthermore, such approaches carry a high degree of risk.

Chapter 5 looked at how the 'need' for a new kind of public and educational management had evolved, in particular how this led to the introduction of KPIs. Specific reference was made to Hood's analysis of the 'mega trends' driving NPM. Callaghan's influential Ruskin speech was an important point in the development of 'new educational management'; although it was the 1988 Education act that was to put in place much of the current day organisational and accountability systems, of which KPIs are a vital and integral part. Shortly after this the citizens' charter was introduced which required schools to set and publish targets as well as publishing test and exam performance.

The rise of testing and assessment was discussed, including the infamous earlier form of 'payment by results'; the revised code. Conceptually, there a number of parallels between this and the underlying philosophy of the new educational management, with both being driven by a perceived need for the education system to be internationally more competitive. The work of the Task Group on Assessment and Testing (TGAT) and the debate on whether testing should be primarily for formative or summative purposes was discussed. Finally, a brief review of the present day curricula and testing arrangements in England and Scotland was carried out. At the secondary level the two systems are quite similar, however for primaries there are significant differences, particularly in terms of the testing arrangements and publication of the results (league tables).

The relatively long and rich history of research on the behavioural effects of PIs in business accounting was considered in chapter 6. This mainly dates from the 1950s: prior to this accounting was viewed as a relatively technical process, and one which did not have inherent behavioural implications. Work from three of the key theorists was considered; firstly Chris Argyris who looked at the effects of budget (target setting) systems and concluded that there was a need for meaningful participation at all levels. Secondly, Anthony Hopwood's work on different supervisory styles, which concluded that a budget constrained style could lead to dysfunctional behaviour, such as the manipulation of the results. Finally, David Otley who takes a broader view and suggests that the actual behaviour is more dependant on the underlying organisational context, rather than necessarily being determined by the technical aspects of the particular accounting system.

A case study on the use of KPI league tables by a distributed organisation (TNT) was also discussed in this chapter. In this organisation performance is compared on a depot by depot basis, and a number of similarities were found between this and educational KPIs which compare schools. Consideration was also given to experimental research on the organisational effects of league tables in business. Work by Keasey *et al* (2000) using a hypothetical supermarket scenario, suggested that actions which would raise a branch's league table position took precedence over those which would improve overall performance.

At a more fundamental level the questions posited by the work of Johnson and Kaplan (1991) was considered. In their book, *Relevance Lost*, they question whether accounting systems are measuring the right things. Two of the most important KPIs used in business are Profit and Return On Investment, and they argue that these are not the most important measures for long term performance; furthermore they argue that they encourage dysfunctional behaviour.

There are many parallels between schools experiences of KPIs and the experiences of those from other areas of the public sector; ultimately they are all 'managed' by the same body – the government. Chapter 7 started by looking at the work of Peter Smith (1995a) who identified nine potentially dysfunctional effects of publishing performance data in the public sector. Many of these are relevant to education, and were used for the research design. In this chapter specific consideration was also given to the difficulties of measuring performance in the health sector, as well as for the railways and police. All of these organisations are under substantial pressure from the government to improve their KPIs, furthermore they face similar challenges measuring what is important. For each a number of examples of the KPIs encouraging dysfunctional behaviour were discussed.

To help assess the impact of KPIs on schools a specific research project was undertaken. Chapter 8 discussed the design, and it was felt that a predominately quantitative approach which included heads and teachers from English and Scottish primary and secondary schools was most suitable. Chapter 9 detailed some preliminary mainly qualitative research which was undertaken. This was in the form of face to face and telephone interviews, and aimed to identify some of the main issues. Furthermore, it provided a useful opportunity to try out some of the potential questions. From this and the previous theoretical chapters a number of hypotheses were formed.

The methodology for the main questionnaire was discussed in chapter 10. Three different question formats; Likert Scales, Semantic Differential and Point scores were identified as being suitable. In addition the survey would allow 'free text' responses to enable some qualitative data to be obtained. In terms of distribution the questionnaires were sent to a random selection of primary and secondary schools in England. In Scotland, however, such a direct approach was unlikely to achieve a reasonable response rate. Therefore four Education Authorities were contacted who gave their approval and support for the survey. Random schools were then identified in each authority. to which the questionnaires were sent. In both Scotland and England copies were sent to the headteacher and a 'random' teacher. In addition, for English schools a copy was also enclosed for the chair of governors[154].

Overall, the response rate from schools (heads) was about 30% (a total 162 responses were received). This was a little disappointing, although not that different to other research considering similar issues. Nevertheless, the responses that were received did appear to be of a high quality. About 35% used the additional comments box, and 40% offered to take part in follow-up research. Two questions suggested that the sample was reasonably representative. Firstly, there was a good

---

[154] In Scotland no direct equivalent, in terms of role and responsibility exists

spread of 'high' to 'low' performing schools, and secondly, most felt that KPIs were a fair reflection of their overall performance; in effect suggesting they were not inherently against KPIs.

The results were analysed in three parts. The first (chapter 11) provides a descriptive analysis of all of the responses combined. In general schools did appear to be in favour of using KPIs systems, with a majority of respondents feeling that they were beneficial, although there was little evidence of parents using KPIs to hold schools accountable. There was a strong indication that schools found KPIs had a narrowing effect on the curriculum, and increased central control.

As would be expected most respondents felt under pressure to meet their targets. Also most wanted to be involved in the setting process, and there was a strong feeling that they (targets) should consider more aspects of the school process. There was however evidence that targets were simply being imposed from 'above', and from the comments some evidence that this was increasing. Many of the respondents felt that the indicator systems did not take account of their particular circumstances, and encouraged a 'blame' culture.

League tables, not too surprisingly were felt by many to be unfair, with the media not acknowledging the limitations and many parents not being able to properly interpret and contextualise the data. There were also signs that many of the respondents were in favour of using value added systems, although there were concerns about publishing these. In general schools felt that indicator systems which they had some control or ownership of, for example, Durham CEM systems, were most useful. This was followed by Ofsted / HMI reports and then KPIs; Benchmark and Panda data in England were considered least useful.

In chapter 12 various comparisons were made in terms of country, phase and post (head / teacher). Only a limited number of questions (14 in all) were used for this analysis; the first group looked at KPIs and the management process, and the second looked at how they (KPIs) may have dysfunctional effects on the schools. The questions from this later group were combined to form a dysfunction index, which was used in this, and the following chapter.

An analysis comparing English and Scottish primary schools (with and without league tables) turned out to be the most interesting and significant. In terms of KPIs and the management process, the results were broadly similar, although it was surprising to find that Scottish primaries appeared to be under more pressure to meet their targets. However, in terms of the potential dysfunctional effects of KPIs there were substantial differences. English primary schools were far more likely to find that their KPIs were overall more harmful, and that they conflicted with school aims, as well as encouraging them to concentrate on meeting targets at the expense of other important objectives.

They also reported that KPIs had more of a narrowing effect on the curriculum, and encouraged schools to concentrate on borderline children. Finally, English primaries were more likely to report that KPIs encouraged a blame culture and were undermining to the individual respondents. In all of these cases the differences were statistically significant, and combined (dysfunction index) the effect was large (effect size 1.68).

The same analysis was carried out at the secondary level for both countries; however, no significant differences were found for either set of questions. This strongly suggests that the significant differences which were found at the primary level, in terms of the dysfunctional effects of KPIs, were in a very large part due to the main difference, namely the publication of the data in public league tables.

Comparisons were also carried out in terms of the different posts; heads and teachers. Quite surprisingly very few differences were found, although it has to be recognised that this may in part have been due to limitations in the sample. It was not possible to confirm that the teachers had been randomly selected by their heads. Nevertheless, at face value this was a very positive finding suggesting that there was little conflict between heads and teachers in terms of their views and perceptions towards KPI systems.

The final part of the results, chapter 13, looked firstly at a follow up telephone survey to assess if the views had significantly changed over the period of about a year. Secondly, it looked at some of the relationships between the results. A total of 32 respondents were re-asked the 14 questions used in chapter 12. In terms of using KPIs for management purposes there were not many differences, with the only significant item being parents making greater reference to KPIs. The reason for this may have been due to the timing, with the second survey being carried out around the time when the league tables were published.

In terms of the second group of questions, two items stood out. The degree of blame culture being attributed to KPIs and how much the respondents found KPIs undermining. It was tentatively suggested that the reason for this was that the respondents (mainly English primary) may have felt less inclined to admit to a blame culture in their school on the telephone, rather than on a questionnaire. This does of course raise questions about the effect of different survey methods.

In terms of the relationship between the results, as might well be expected, there was a strong positive correlation between a schools relative performance, and how fair they felt the indicators to be. However, more surprisingly, schools which performed well, were also more likely to report dysfunctional behaviour. Not so surprisingly though, was the finding that parents with children at

'high' performing schools were more likely to discuss KPI performance. It therefore follows that parents at 'low' performing schools are less likely to put pressure on 'their' school to improve. which of course was one of the main reasons for having KPIs and league tables in the first place.

Five key themes

This section brings together some of the issues raised in the theoretical and research sections, by looking more specifically at five key themes which run throughout the thesis.

Accountability

This is very much at the heart of the new educational management. KPIs play an important part in the accountability process, both directly and indirectly (eg. Ofsted). The government's rationale was and still very much is that increased accountability will lead to higher standards which will be of benefit to the children receiving education and the wider community. As a general principle the notion of holding organisations and individuals accountable is of course quite reasonable. However, to believe that this will automatically lead to better performance, both in terms of what is actually done, and meeting broader organisational aims, may be naïve; and the process itself may lead to dysfunctional behaviour. For example, the high levels of 'blame culture' which were evident in the findings may well in part be due to the accountability systems.

The government's espoused view is that accountability should be both ways: 'up the line' to the government, and 'down the line' to parents, students and other local stakeholders. There was much evidence of the former, with schools feeling accountable for their KPI targets, and through this accountability chain, the DfES being accountable to the Treasury for their PSA targets. However, in spite of various claims and explicit policy (eg. Citizens /parents charter) there was little evidence of the latter, and there are perhaps arguments that this 'down the line' accountability has if anything, decreased.

In some instances it may well be that schools are less accountable for issues which parents and the local community feel are important. (Evidence that KPIs may not be that important to parents was discussed in chapter 11). Where the wishes of the various stakeholders do not coincide, it would seem likely that accountability based on meeting KPI targets will dominate. Furthermore, it will be very difficult for a parent to question or challenge a school which is deemed to be performing well, in terms of KPIs and Ofsted (the two generally go hand in hand). And in this situation the school may well be oversubscribed and the parent's ultimate sanction will be relatively meaningless.

Moreover, accountability to the locally elected councils is now to all intents and purposes non-existent. In England the LEAs role and responsibilities have been substantially curtailed, and unless there are serious problems they now have very little influence on the organisation and direction of the school. This largely contrasts to the relation between schools and EAs in Scotland (IES Scotland 2000), although in reality this varies between authorities.

Professional accountability may also be in conflict with 'up the line' KPI accountability. There was evidence in the survey of heads and teachers being forced to meet KPI targets, which conflicted with their professional judgements. For example, rigidly following the literacy strategy rather then teaching what they felt most appropriate. The 'needs' of the school to get as many level 4's may not coincide with the *needs* of the children. Furthermore, there is increasing evidence (Chrisafis 2002) of teachers' employment contracts only being renewed if KPI targets are met, and this can only exacerbate any such problem.

Therefore, although some degree of accountability is quite right and proper; and if carefully incorporated in to effective management systems, will doubtlessly be of benefit (eg. providing feedback), it would seem wrong to simply believe that schools and those working in them are necessarily more accountable to the right people for the right things. In essence subtle complex organisations, such as schools, need subtle complex accountability systems, which are based more on trust and professionalism, rather than fear and mechanical processes.

Targets

An important part of the accountability process is using and setting targets. Like accountability systems they have tended to be used in education in crude and simplistic ways, and with little regard to the more subtle and complex features of the process. This has been very much the case throughout the public sector, where if a problem is identified (often by the media) a target is slapped on, and 'magically' the problem goes away[155]. The evidence from the business accounting sector which was discussed in chapter 6 illustrated the importance of taking both a long term view, (change is a slow process), and considering the wider behavioural consequences.

Even so, the research here found that many of the respondents valued targets and target setting, and used these to inform their work and practice. Most would agree that targets and targets setting are important components in effective teaching; for example, children like to measure themselves against targets, whether in a football match or a spelling test. And for teachers targets provide an important basis for constructive feedback to their pupils, and heads to teachers and so on. The

---

[155] Specific crime reduction targets would be a good example

problem would seem to be the miss-use of targets. for example, arbitrarily imposing targets or using targets as a basis of blame and retribution.

The importance of participation and sharing ownership of targets was discussed in chapter 6, and the evidence from the survey was that virtually all of the respondents wanted to take part in this process. However, there was evidence in the comments that targets were simply and arbitrarily imposed on schools, indeed there are indications that centrally imposed targets will increase: not least to 'help' meet the 2004 / 6 PSA targets.

Although, as stated above there was evidence that schools did value targets, the research did not consider in too much detail why this was so. Optimistically, it may be that they are used primarily to improve the overall quality of the teaching. However, it may also be that in some circumstances they are used primarily to help meet the school's overall targets: in other words they have become an end in themselves. Or put another way, to meet the school's needs rather than those of the children. It can perhaps to an extent be argued that a school viewed as 'successful' is of benefit to the children and local community, although this might have a negative impact elsewhere in the system.

Complexity

The issue of complexity is very pertinent to the paradoxical nature of the 'new educational management'. At the policy level there is an apparent lack of recognition or perhaps even denial of the inherent complexities of the education process. In chapter 4, rational planning systems largely originating from the work of Weber and Fayol, were discussed: this rational approach very much runs throughout educational policy formulation and implementation, with the belief that problems can simply be identified rationally, and then solutions logically prescribed.

To a large degree it would appear that the problem is not simply one of lack of imagination or intellectual rigour on the part of those responsible for policy, but rather a system driven by the need for quick (and decisive) fixes and irrefutable evidence of progress. In this context, picking a limited number of isolated issues 'to fix', is preferable to broader more systemic issues. Another advantage of simplicity, as opposed to complexity, is that it very much equates to transparency, rather than opacity. And therefore simple clear messages can be 'sold' to the public, rather than having to remain in the domain of the 'so called' experts. However, many of the important less obvious and more subtle issues may well be lost with this approach. Furthermore, it makes joined up approaches to policy and practice less attractive. It is easier for the education system as a whole to concentrate on exam KPIs, rather than giving too much consideration or resourcing to important cross cutting targets: for example, the sure start PSAs or local targets to reduce teenage pregnancy.

At one level the government's reasoning appears to recognise the subtleties and complexities of organisational management, it is however the application that too often seems to fall short. For example, the Performance and Innovation Unit's (2001) discussion paper, *Better Policy Delivery and Design*, (NB. *delivery* before *design* might be telling!), provides an excellent analysis and arguments on how policy 'should' be formulated. Indeed, many of the points made in this thesis would be at home in this document, and vice versa. For instance, the dangers of dysfunctional behaviour from 'high stakes' indicators, and the problems associated with a 'blame culture' are recognised and discussed, as is the need for a learning organisational structure. However, the difficulties appear to begin when this thinking is related to actual practice. In the case of this particular paper they use the implementation of the Literacy Hour as a model of good practice. And no amount of bending of their theory manages to square it with their description of the implementation strategy and the quoted Ofsted evaluation.

The right measures

However accurate and all encompassing the indicator, the relative performance of complex organisations or systems can not be reduced to one or even several measures. A range of indicators are needed to effectively monitor performance, and even a large set of indicators needs to be supplemented by other indicators (which may vary depending on the circumstances). Furthermore, there needs to be the facility to 'drill' down to the underlying factors; for example, aggregate class indicators are of little use in setting appropriate targets and work for individual students; the same is true at the School and L(EA) levels.

Nevertheless, using a large number of indicators can also present problems. For example, quite naturally people can only comprehend and appreciate several indicators at a time, and trying to read too much in to a large set of indicators may lead to 'paralysis by analysis'. Therefore, when looking at the question of what are the 'right measures' it is important to consider who, and what, are they for. As they stand, KPIs with for example the addition of confidence intervals, may show extremes of performance and therefore may have a use as early warning devices, but it is futile to expect much more. From this perspective, later in this chapter the case is made for using a balanced set of indicators.

Broader Organisational Context

At the heart of many of the issues considered in this thesis is the effect of the broader organisational context. Different schools and perspectives of organisational theory were discussed in chapter 4, and it was suggested that the organisation and management of the education systems was based more on a 'rational / scientific' model, rather than a 'learning organisation' model. The

research element of the thesis very much supported this notion. PIs play an important part in the management and control of individuals in this rational model, however this approach can easily discourage learning. For example, performance which is deemed to be sub-standard (or just a different standard) will be viewed as unsatisfactory and the appropriate correction must be applied; rather than allowing both the individual and the broader system to learn and develop from the experience. Whilst it is of course not being suggested that children should receive a sub-standard education delivered by incompetent teachers, the best and most appropriate methods are rarely applicable to all situations or for all students. Practitioners need to be given the confidence, and trusted to use and apply the available evidence to inform their practice, as well as contributing their experience to the overall body of knowledge.

An important part of the learning organisation approach is the issue of how individuals are managed. One arguably dysfunctional consequence of the NPM principles has been the reduction in professional autonomy, for example the blanket imposition of national strategies. The work of McGregor and his theory X and Y was discussed in chapter 4, and the evidence from this thesis strongly supports the notion that individuals are currently managed, and perhaps increasingly so, in terms of theory X. This theory takes a relatively pessimistic view by assuming that targets need to be set, few alternatives should be available, there should be clear direct lines of responsibility and appraisal should focus on past performance.

There would seem little doubt that a learning organisation approach to the management and organisation of the education system is the most appropriate way forward, indeed this is the method advocated by many government publications (see next section). Although, as Nutley and Davies (2000a) point out, this approach may encourage schools and individuals to develop in many different ways, and by definition force the government to relinquish some degree of control. It should be noted that it is the English government's espoused policy that schools should be able to 'earn' their freedom to innovate; however the criteria (or criterion) for allowing this will be a high and sustained level of KPI performance. In other words schools will be able to be more innovative at meeting their targets – however, as argued previously, this would not seem a very effective way of developing the education system as a whole, and ultimately, in being able to meet the increasingly complex future needs of the country.

The future of KPIs

Since the survey was carried out there have been a number of changes to the KPIs, in particular how they are reported. There have also been on a number of occasions problems with the league tables. In Scotland the tables were not produced in 2000 because of difficulties with the exam

results. They were published in 2001 but with many anomalies, caveats and complex explanations – in effect making them difficult to properly interpret. In England widespread concerns over the A level gradings in 2002, has raised questions over whether these results will be included in the tables, which are due out in mid December 2002.

The most significant change to the English secondary tables will be the inclusion of Key Stage 3 tests, and a 'progress indicator' (value added)[156] will be provided between KS2 and KS3. It will be interesting to see what effect this will have on English secondary schools, particularly lower schools. A similar 'progress indicator' for primary schools between KS1 and KS2 has been postponed until 2003. In other areas of the UK there have been important changes. Performance tables are no longer published in Wales and Northern Ireland, with schools being responsible for disseminating their own results. Changes have also occurred in terms of the actual tests in Wales, and in Northern Ireland it looks as if the 11plus will be abolished, although much depends on the future of the assembly.

In terms of targets, the PSAs on which educational KPIs are based have been 'cranked up' from the 2002 levels (Treasury 2000). And as promised Estelle Morris resigned as Education secretary, having given a pledge that the 2002 targets would be met – a sad move[157] which sends out all the wrong messages about targets.

There is little doubt that PIs and KPIs will continue to play an important part in the management of the education system. However, there would appear to be some confusion as to how the system will develop in the longer term. Talbot (2000) points to the conflicting government policy, with on the one hand 'hard' output orientated systems, ie. the current KPIs, and on the other, 'soft' process and outcome based systems, ie. balanced packages. The eventual system and its underlying philosophy will substantially influence the effect of PIs on schools, with it is argued a 'balanced system' being a necessary pre-requisite for the development of a 'learning organisation'.

The white paper, *Modernising Government* (Cabinet Office 1999) advocates a 'balanced scorecard' system based on the European Foundation for Quality Management (EFQM) Business Excellence Model. This approach, which is used by various private sector organisations, attempts to avoid many of the pitfalls of single high stakes output indicators, by using a wide range of indicators, including process and outcome. There is much support for such systems in the public sector (Woodcock 1998, Gambles 1999). The EFQM system has been proposed for use in Scottish

---

[156] It should be noted that these progress indicators are very crude and potentially misleading. Even the DfES described them as "not ideal". (TES 2002)
[157] Other factors may have contributed to this decision

schools (Audit Unit 1999b), and the Durham CEM systems also provide a similar range of indicators.

However, the Treasury's Public Services Productivity Panel report (Mayo 2000) proposes a very different, (but rather familiar), system for education. This is based on the PSA targets, (ie. the current KPIs), with the addition of a relative value for money indicator, which would, for example, show the cost in financial terms of a GCSE. Similar figures in the ubiquitous league table format have already been produced for independent schools (Sunday Times 2001).

There is little doubt that these two approaches would have very different effects on the culture of the education system, which would in turn influence 'organisational learning', as well as the place of evidence in policy making. At present it would appear that the second 'harder' (KPI based) approach will prevail. The recent white paper, *Schools Achieving Success* (DfES 2001a), clearly equates success with performance measured by KPIs, and makes no mention of balanced scorecards. The paper does however state that successful schools will be allowed greater freedom, although within, and controlled by, the current KPI framework; "Where schools are successful...The framework of performance targets and accountability...must remain in place" (p.42). As argued in the previous section this would seem to be a recipe for maintaining the status quo rather the providing the opportunity to develop useful and effective indicator systems.

At a national level, two other factors would seem to point to the continuation of the current system. Firstly, as already mentioned, looming over the DfES are the new PSA targets for 2004 and 2006. Given the political importance of education and the proximity of the next election, the government will want to keep everyone's 'eye on the ball', to help achieve these new targets. Therefore, even if a national 'balanced scorecard' system is adopted, high stakes KPIs based on the PSA targets may well still, in reality, prevail. And although the new Education Secretary has explicitly refused to give any undertakings, there will still be great pressure on him personally to meet PSA / KPI targets.

Secondly, central to the government's approach on the development of the public sector is the increasing involvement of private sector organisations (DfES 2001d). Measuring performance is an important part in this process, and 'hard' unambiguous indicators are far more suited, and indeed essential, for the necessary contractual arrangements. For example, the contract awarded to Cambridge Education Associates (CEA) to run Islington education authority set targets in terms of GCSE passes. As these were not met CEA was 'fined' (Garner 2001); it would however have been far more difficult or impractical to write similar penalty clauses into a contract based on a 'balanced score card'.

Therefore, even though at the ideological and philosophical level there may be evidence that the current government would like to adopt a 'learning organisation' approach, which uses a balanced range of indicators, such an approach seems a long way off. Furthermore, at a more pragmatic level, 'hard' KPI systems which largely reject the notion of organisational learning, appear to produce the results, and have allowed 'objective' claims to be made that the quality of the service has improved. This latter approach has many political benefits, however whether it supports the realization of the many complex long-term aims of the education system is another matter; as is whether standards have in reality risen by as much as is claimed. At the primary level this is very much open to question in some areas of the curriculum, (Tymms and Fitz-Gibbon 2001, Gold 2002). And at the other end, the report from the Engineering Council (2000) *Measuring the Mathematics Problem,* gives rise to concern. This paints a rather bleak picture of the standards of mathematics, and gives a very different impression to that given by the apparent substantial rises in 'A' level performance over the last few years.

Furthermore, there is evidence that parents too are unconvinced about the rises in educational performance (Passmore 2002). As well as increasing evidence that certain groups of children are not progressing as they should. This was discussed in a recent systematic review[158] of the evidence on the impact of summative assessment and tests on students' motivation (Harlen and Deakin Crick). Key findings suggested that high stakes summative assessments, have a negative effect on the esteem of lower achieving students, and that there is a widening[159] of the performance gap[160]. They also help create an environment which favours a particular type of pupil/style of learning, and that this type of learning may not best equip young people to 'solve' the problems of the future. Significant to this thesis, they make a number of policy recommendations including; improving the quality of information systems in schools, discontinuing league tables[161], avoid targets based on high stakes summative assessments, and interestingly from the findings presented here in terms of English and Scottish primary schools;

> For summative purposes in reporting on individual students, move towards testing students when their teachers judge them to be ready to show their achievement at a certain level, as in the Scottish system for national testing in the 5–14 programme, thus avoiding experience of failure and its impact on self-esteem.
>
> (p. 78)

Taken together the issues raised in this section illustrate many of the challenges and problems facing policy makers in terms of the future of KPI systems. Whilst the government may well want

---

[158] EPPI
[159] Interestingly at the school level, Gorard (2000) found that the gap between schools was reducing
[160] League tables are too crude a measure to show this kind of pupil level data
[161] The phrase 'league tables' wasn't used but rather; "avoiding comparisons among schools in terms of test results..."

to continue to use high stakes KPIs to control and hold schools and individuals accountable, as well as other stakeholders (eg. PFI contractors), there would seem little evidence that they are effective in measuring what is important or indeed what many people want.

Critique

There were four main elements to the thesis; the first, the broader organisational perspective discussed how PIs are a part of, and interact with other organisational and management issues. In particular the influence of different structural models was considered mainly with respect to generic management theory. This approach was found to be very useful in developing an understanding the organisational impact of KPIs on schools. In essence the effect of KPIs was found to be very much determined by the broader organisational values and contexts. For example, KPIs can in one situation be used to help and support individuals and encourage the organisation to learn, on the other hand they may be used as a basis of blame and retribution, and this will discourage organisational learning. The evidence from the theory and the research in this thesis, was that the latter model in reality prevails, in spite of claims to the contrary (and indeed probably a genuine desire on the part of policy makers). It was suggested that although many educational theorists have tended to avoid generic management models, which have their origins in the industrial sector, they do nevertheless provide many useful insights and conceptual models, which are as relevant to schools and the broader education system as they are to other complex organisations.

Similar conclusions were reached with respect to business accounting theory. Although the terminology and many of the specific measures are different; in essence, both business and education systems are attempting to capture and represent the efforts and creativity, and ultimately the contributions made by individuals within the broader system. And as in education the headline figures are rarely a true indication of performance, and indeed may in themselves be a distraction or a dysfunctional influence. Therefore, it was felt that much can be learnt from the business sector, in particular how broad based balanced indicator systems may be used effectively. It was also evident that this learning process should be both ways. There is much that the business sector can learn from the experiences of the education sector and their PI systems.

Likewise, it was found that many of the experiences from other parts of the public sector were very relevant to the education system; ultimately 'we' are all managed by the same body. And there is evidence that the performance measurement and accountability systems, (eg. PSAs and targets), have similar behavioural and organisational effects. There is therefore much that can be learnt from

each other, with the further potential benefit that this may help facilitate joined up thinking and working.

The actual research broadly supported much of the theory from the earlier chapters, however it also provided some interesting and unexpected results. For example, the substantial differences between English and Scottish primary schools, as well as this, it helped identify a number of areas worthy of more detailed research. In terms of the main survey a number of important limitations need to be acknowledged. The most significant was the response rate, which was a little disappointing. It is difficult to draw any firm conclusions about the reasons for the non – responses, although it did seem that the most likely reason was simply the time constraints on schools. Nevertheless, the views of those non-responders may have been significant, and it is possible they might have influenced the results to some degree. Retrospectively, strategies at the initial design phase could have increased the response rate; for example, Thietart *et al* (2001) suggest that improvements of 20 – 30% can be achieved with appropriate follow up strategies[162]. Furthermore, the response rate from the teachers was disappointing, although as pointed out in chapter 10, this appears to bedevil much educational research. Nevertheless, it was felt that there were short-comings in the design of the teacher survey, in particular the potential bias from heads being involved in the distribution, and again in retrospect a more direct approach could have been used.

However, at a more fundamental level the use of a mainly quantitative survey may have missed or even hidden some important factors. By definition a questionnaire has to mainly consider the issues which have previously been identified as important and worthy of assessment, (although some additional factors can be included in the 'free text' part). As discussed in chapter 8, there is something of a paradox in terms of research design, with qualitative research taking a broader view and hoping to provide the questions, and then quantitative research taking a narrower view, and hoping to provide the answers. In terms of this thesis it was felt that some of the key issues or questions still remain unidentified. Therefore, further research would benefit from both more qualitative, and then more specific experimental type quantitative research; in effect a cycle of qualitative and quantitative. That said, there are currently a number of areas worthy of detailed experimental research. For example, potential or forthcoming changes to the indicator systems, such as the inclusion of KS3 results and 'value added' in England, baseline assessments in Scotland, and the abandonment of league tables in Wales and Northern Ireland.

One further limitation of the research, and in particular the impact of its conclusions, has been due to the lack of relevant comparative research. Five years ago it had been expected that there would

---

[162] It is of course the case that schools and heads are probably something of an atypical group in terms of surveys, in that they are perhaps less likely to respond to follow up strategies

be a number of other similar research projects; however this has not really been the case. During this period there have been many observational types of research; but few have provided much indication of the extent to which KPIs have an effect on schools[163]. Probably the most relevant recent research was the NUT survey by Neill (2002)[164]. This looked at a large sample of teachers[165] (>3000 NUT members) and for a number of the items there was broad agreement with the findings presented here. For example, 86% of the respondents found that the Key Stage tests had a narrowing effect on the curriculum, and 87% were against league tables. One interesting finding was that only about 8% of teachers felt that the tests improved children's motivation, and even less (6%) said the test had improved or increased parental interest in their child's education. Whilst these findings do make interesting reading, it is difficult to make any significant valid comparisons with the work carried out here; furthermore, there may be potential sample bias, although in fairness this was discussed at length by Neill.

There are however increasing calls from other bodies for more research and the proper evaluation of educational interventions, for example, the Audit Commission; and more specifically for the issues considered by this thesis, the independent evaluation (by OISEUT) of the NLNS:

> DfES might wish to evaluate, on an ongoing basis, the intended and unintended effects of the use of national targets and high stakes Key Stage 2 tests. This would include, for instance, a much more systematic examination than we can undertake of the extent to which teachers in Year 6 focus on test preparation and teaching to the test, and what this really means. It might also include, for a random group of schools, the use of somewhat different tests for reading, writing and mathematics, to see if the gains hold up or if they are restricted to a particular type of test.
>
> (Earl *et al* 2001: 83)

As well as this, interest in educational KPIs is increasingly being shown by other sectors. For example, Telhaj (2002) from the Economics department of Staffordshire University, presented a paper at BERA looking at the dysfunctional effects of PIs in English education, using a Principle – Agent framework. The work is on-going[166] and very much builds on previous work of Adnett (1999 and 2001) at the same institution. Again a number of findings resonate with those presented here, for example, the focusing on C/D border line pupils and 'cream-skimming'. Again this very much supports the arguments for increased inter-organisational learning.

In terms of supporting research on issues such as KPIs, it is recognised that this is almost certainly bound to be inherently critical of government policy (how else does a complex system learn?) And

---

[163] For example, West and Pennell (2000) make a number of interesting observations, but avoids the important and more controversial underlying issues

[164] Research for ATL (2001) a year before drew similar conclusions (see also Wiliam 2002 author of ATL report)

[165] Of which 2% were heads

[166] PhD study

as Davies (1999)[167] points out there is an increasing trend for the government to support (and then 'pick') research and evidence which supports the policy first thought of. Furthermore, as Cowie (2000) observes, with equal relevance to Scotland and England, 'we' are not encouraged to ask the really important questions: "Professionals in Scotland are not encouraged to discuss fundamental questions of meaning and purpose." (p25).

End Note

The majority of the work for this thesis was carried out between 1997 and 2002. This has been very much a period of consolidation of the principles enshrined in the 1988 Education Reform Act, as well as the broader philosophy of 'New Public Management'. Although many of the claims have been to the contrary, the effect has been to move the direction and control to the centre. Associated with this has been the development of a pernicious form of accountability and one to which KPIs are inextricably linked.

The vision of a true learning organisational structure is still a long way off, indeed during this period it may have even moved further away. In England, the practice of holding schools to account by publishing raw ranked performance data and 'naming and shaming', looks set to continue. Comments by the latest Education secretary (in October 2002): "The [KS] targets...are absolutely critical to everything we are about" (Ward *et al* 2002)[168], does little to instil confidence. Moreover, in England the use of KPIs and league tables is increasing in other related areas. Local Authority league tables, which are based on the comprehensive performance assessment inspections (due mid December 2002), will doubtlessly have an impact on their relationship with schools, as well as for the special needs and looked after children, for whom they have a direct educational responsibility.

Perhaps, the 'best hope' for progress towards a more enlightened approach to the direction and management of the education system, lies in learning with, and from, the other countries which make up the United Kingdom, as well as perhaps those in Europe and the rest of the world. In terms of exam performance: Scotland, Wales and Northern Ireland, with their increasingly different and nationally distinctive approaches are still ahead[169].

Whilst technical improvements can, and perhaps are being made to the KPI systems, this is only one side of the coin. Such improvements need to be accompanied by, and to be a part of, more

---

[167] And many others (eg. Ball 1998, Wallace 2001, Davies *et al* (1999)

[168] Indeed echoes of his conservative name sake and previous education minister

[169] England still lags behind the rest of UK (Woodward 2002b)

fundamental changes to the underlying philosophy and organisation of the education system. And such a change requires 'top down' leadership: one which devolves trust and responsibility, rather than blame and accountability. However, very much working against this, is the inherent dysfunctional adversarial political processes, which often force government and opposition in to 'diametrically illogical' positions; and from which only prejudicial rhetoric appears to emanate. Recent attacks by the current opposition[170] on the 'failure' of the government to meet 75% of the PSA targets, (as well as a ministerial scalp), illustrates the problem. This generates a climate in which it is difficult for a government, whatever its persuasions and aspirations, to create and nurture a culture which both encourages and learns from debate and evidence - whatever the source. It is this macro political culture which must change if we are to realise the improvements in education desired by all.

[170] Howard 2000

Appendices

Appendix 1 Letter sent confirming meeting for preliminary research. (1 page)

Appendix 2 Questionnaire used to 'guide' preliminary research meeting, completed by author (3 pages)

Appendix 3 Feedback form for main questionnaire, completed by those doing trial (1 page)

Appendix 4 Letter sent with the main questionnaire packs (1 page)

Appendix 5 The basic questionnaire sent to English primary and secondary schools. Slight alterations were made for the Scottish schools (6 pages)

Appendix 6 Feedback letter Scotland. Slight alterations for English respondents (1 page)

Appendix 7 Follow-up research telephone questionnaire. Slight changes were made in the terminology for the different schools (1 page)

The Rowans
Catton
Thirsk
North Yorkshire
YO7 4SH

Mr ABC
ABC School
ABC Rd
N YORKSHIRE
HG1 1AB

20th March 1998

Dear Mr ABC

Preliminary Survey – Impact of Key Performance Indicators

Thank you for agreeing to meet with me on the 25th March 1998.

In advance of this meeting I thought it might be helpful to provide you with some further background information, and details of what I hope to gain.

My starting point for this research is to look at the effect of Performance Indicators in particular, Key Performance Indicators (KPIs), on the Organisation and Management of secondary and primary schools in England and Scotland. I am defining KPIs as those indicators which are made public; in effect, Test and Exam performance as well attendance figures. The most relevant time period under consideration is the last 15 or so years.

Broadly speaking, my main aim will be to understand your indicator system and how it affects your school, in particular, the organisation and management processes. In addition, I hope to identify issues, which will be of relevance to a future larger scale survey.

I would like to start with a brief discussion about your school's indicator systems, and some more general contextual issues. Then to move on to questions in the following areas:

- **Internal indicator systems** - indicators used by the school for your own management and development purposes.
- **External (public) indicator systems** - eg, league tables, OFSTED action plans, and national target setting.
- **The use and effects of the indicator systems** - in particular how they influence behaviour, and relate to your strategic management.
- **Developments** - Looking ahead to the future of indicator systems.

I look forward to meeting with you on Wednesday.

Andy Wiggins

Name of School

LEA

Type

Size

3rd party indicator systems?

Do you find it useful.....

Main strategic Aims of school

Key success factors

Brief Description of main internal indicator system

What are the main Internal Indicators you use for day to day management

What are the main Internal Indicators for strategic management.

Are they generally well understood and used by

Staff

Governors

Is a lot of work and effort required for their collection / compilation

External Indicators - Public, league tables, OFSTED reports:

Do the external indicators measure the most important aspects of your school

Do they provide a valid assessment of your School

- Valid assessment of other local schools
- Valid assessment of education system as a whole

Are they properly / appropriately understood by the:

Governor

Current Parents

Seven point scale : Strongly Yes - 7 to Strongly No - 1 Unable to answer - 0

Public in general

Are they important for your short term planning / management?

Are they important are they for your long term / strategic planning - Development Plan eg (targets)?

Do the league tables help improve the quality of education in your school?

Do the league tables help improve the quality of the education system as a whole?

Does the competition element of league tables help raise standards?

Does increased public accountability raise standards?
Will the inclusion of a Value Added figure improve the value of the tables?

Is there need for more indicators to be made public and included in the tables

Free school meals are sometimes used as a means of comparing similar schools - is this valid

The use and effects of Indicator systems

Do your internal and external systems give a broadly similar picture, in terms of performance

Does your indicator system provide a balanced picture of your school

Would this be the view of the other Staff

" " Governors

" " Parents

Does your indicator system help identify your key success factors

Does your indicator system help achieve your strategic objectives

Does you indicator system help identify your schools weaknesses

Do the external indicators tend to 'override' or take preference over internal indicators

Do the external indicators tend to 'override' professional judgements

Do you feel your indicator systems give a reasonable indication of current performance

Do you feel your indicator systems give a reasonable indication of future performance

Do they discourage your school from pursuing educational objectives which you (the school) feel important, but which are not measured?

Do they encourage the school to concentrate on meeting short term targets at the possible expense of longer term objectives?

Seven point scale : Strongly Yes - 7 to Strongly No - 1 Unable to answer - 0

Is this likely to increase with the new development such as national target setting.

Do they encourage the school to concentrate on specific pupils. such as border line grade 'D students?

Do they discourage other courses which may be more appropriate for some students?

Do they encourage competition between teachers?

Is this good constructive competition

Do they encourage competition between departments?

Is this good constructive competition

Do they encourage competition between other local schools?

Is this good constructive competition

Do they encourage 'gaming' behaviour. for example setting low targets which are easily achieved?

Will the current developments in target setting encourage this

Do they encourage or discourage innovation at the class / Teacher level?

Do they encourage or discourage innovation at the School level

Are they used effectively by the LEA in their decision making and planning processes?

Developments

Is the principal of linking teacher appraisal to performance related pay good

Is the principal of linking school performance to the level of funds it receives good

Overall do you feel that the national target setting - valid / beneficial to schools

Should targets be applied on a 'national' basis as opposed to negotiated individually

Are there any other areas of the overall school management process which you feel would benefit from having indicators available for

Are there areas that we should not try to measure / quantify

Finally, do the external indicators and associated league table provide 'Value for money'?

Seven point scale : Strongly Yes - 7 to Strongly No - 1 Unable to answer - 0

University of Durham

Please reply to:
The Rowans
Catton
Thirsk
YO7 4SH
Tel: 01845 577760 or
andy.wiggins@durham.ac.uk

Questionnaire Evaluation Sheet

Thank you very much for agreeing to help trial this questionnaire.

Please complete the questionnaire as you would normally. Then if you feel appropriate mark (in red of course!) any mistakes, unclear questions or explanations etc. Finally, if possible please complete the following sections:

What was your first impression, and is it likely that you would have gone on to complete it	
About how long did it take, and did you feel the length was about right	
What did you think about the layout	
Was the introduction clear enough	
How was the overall 'tone' of the questionnaire	
Overall, did you find it interesting	
How about the level - was it over-technical or too simplistic	
How did you feel about the Head's introductory letter	

If you wish, please make any further comments on the back of this form, and add any other questions which you feel could / should be asked.

Again many thanks, and I will send you a copy of the summary of the results in due course.

Andy Wiggins

University of Durham

PPI Project
Freepost NEA242
Durham
DH1 1BR

FAO: The Head

The Effect of Key Performance Indicators on Schools

Three perspectives

Governors - Heads - Teachers

I am carrying out research into the effects of Key Performance Indicators as part of a PhD. Your school has been picked at random, (as part of a sample of three hundred nationally), and I would be very grateful for your help with a questionnaire.

Introduction: The use and significance of Key Performance Indicators (KPIs), such as Key Stage tests and GCSE results, has of course greatly increased over the years. This research aims to improve our understanding of their effect on the management and organisation of schools.

Organisation: I am considering the question from three perspectives; that of Governors, Heads, and Teachers. Accordingly, enclosed are three copies of the survey; for yourself, the Chair of Governors and a teacher at your school. I would be very grateful if you could distribute these, (attached to each is a freepost return envelope). The questionnaires take about 15 minutes to complete.

For the validity of the research it is necessary to randomise (as far as possible) the selection of a teacher. Would you please pass it on to the teacher with the surname beginning with M, (or the next letter after this in the alphabet) with at least 2 years teaching experience, and some knowledge and experience of PPIs.

Confidentiality: The contents of the questionnaires will only be known to myself, and will remain confidential. The questionnaires are numbered to analyse the return pattern; please delete these if you wish your school to remain entirely anonymous.

Finally: I am very conscious of the time constraints on both you and your colleagues, and would very much hope that you feel this is an important and worthwhile area of research. I will be producing a short summary report on the main findings of this survey, and would be delighted to send you a copy. (Please complete the details at the end of the questionnaire). Thank you.

Should you have any queries do please contact me
Tel 01845 577760 or andy.wiggins@durham.ac.uk

Andy Wiggins - May 1999

University of Durham

Ref.....

The effect of Key Performance Indicators on Schools

Questionnaire

Please return in the Freepost envelope by .....

Introduction:

Thank you very much for taking part in this survey, which looks at the effects of Key Performance Indicators (KPIs) on the management and organisation of schools, and is part of my PhD.

The significance of KPIs in education has dramatically increased over the last few years. this research aims to contribute to our understanding of how they affect schools. For the purpose of this survey it might be helpful to define what I mean by KPIs.

KPIs are the performance indicators which *must* be made public. Currently, this includes:-

- Key stage results
- External examination results
- The number of children with special educational needs
- Pupil attendance data
- Statutory KS2 and GCSE targets

Confidentiality: Your answers to this questionnaire will of course remain confidential. If you prefer not to be identified please do not add the name of your school. The questionnaires are numbered (top right cover page) to analyse the return pattern, please delete this number if you wish to remain entirely anonymous.

The Results: When I have completed this survey I would be delighted to send you a summary of the results; please add your details to the end of the last page.

The Questions:

- Please complete as many of the questions as possible.
- This questionnaire will be used by Governors, Heads and Teachers from both primary and secondary schools; therefore some of the questions may not appear to be relevant to you - *Nevertheless*, your responses, from your perspective, would be very useful.
- If there are still questions which you cannot answer, please mark them *Don't Know* (**d/k**).
- If you wish to make any further comments please feel free to do so on the form, or on the back page.

To help put your answers into context please complete the following

A few details about you and your school

For office use

Your Position: ***Governor / Head / Teacher***

Your school: .....

How long have you been in your position:.....

Which type of school is this: ***Primary / Secondary***

If applicable: Which subject / year group do you teach:

.....

Which LEA are you in: .....

How would you assess your schools position in the LEA league tables?

High	
Quite High	
Middle	
Quite Low	
Low	

Is this a fair reflection of your school's overall performance?

Please score out of five.
1 very unfair to 5 very fair

Please *tick* any indicator system which your school uses:

QUASE	
PIPS	
MidYIS	
YELLIS	
ALIS	

Please give details: .....

The use and people's awareness of KPIs

	YES		Neither or	NO	
	Definitely	Mainly	Yes & No	Mainly	Definitely
Are you able to recall your school's approximate figures from the performance tables					
In your routine dealings with parents do they frequently refer to the performance tables					
In your dealings with prospective parents do they frequently refer to the performance tables					
Are your school aims mostly expressed in KPI terms					
Is your school's position in the league tables very important to you, personally					

b

*Each pair of statements represent the extremes, please put an **X** in the position which best reflects your view:*

For example: I hope it's sunny tomorrow ---✗---------- I hope it rains

How do you feel overall about KPIs

c

Beneficial	----------	Harmful
Discourage innovation	----------	Encourage innovation
Broaden the Curriculum	----------	Narrow the curriculum
Undermining	----------	Supportive
Trusted	----------	Mistrusted
Frequently discussed	----------	Rarely discussed
Encourage local control	----------	Encourage central control

Targets and Target setting

	YES		Neither or	NO	
	Definitely	Mainly	Yes & No	Mainly	Definitely
Is there great pressure on your school to meet KPI targets					
Are your school targets mostly expressed in KPI terms					
Is there any conflict between your school aims and targets					
If appropriate, do the targets usually override the aims					
Does your school tend to avoid activities which are not measured by KPIs					
Should targets be set on an individual pupil basis, rather than for the whole school					
Do you feel that you should participate fully in the setting of your targets					

d

Statutory targets - more specifically

	YES		Neither or	NO	
	Definitely	Mainly	Yes & No	Mainly	Definitely
Should the statutory targets consider more aspects of your school					
Will the use of statutory targets help raise overall educational standards in your school					
Do you feel that your school will tend to concentrate on meeting the statutory targets at the expense of other important objectives					
Are you confident that your statutory targets will be achieved					
Do you believe that your school's targets will allow for your particular circumstances					
Do you feel that your school has access to suitable information for setting the targets					
Has the target setting process changed your school's relationship with the LEA					

e

Each pair of statements represent the extremes, please put an X in the position which best reflects your view:

Your feelings on Targets and target setting

A hindrance	----------	Useful
Help raise standards	----------	Don't help
Encourage blame culture	----------	Discourage blame culture
Relevant to all pupils	----------	Relevant to no pupils
Good for pupils	----------	Bad for pupils
Restrictive	----------	Empowering
Improve relationships between staff	----------	Worsen relationships between staff

KPIs and judging performance

	YES		Neither or	NO	
	Definitely	Mainly	Yes & No	Mainly	Definitely
Do the media generally acknowledge the limitations of KPIs					
Do KPIs provide an indication of the most suitable school for a particular child					
In your experience do parents have an understanding of the limitations of KPIs					
Do KPIs have higher credibility in your school than other internal indicators					
Do KPIs have a higher credibility in your school than 'professional' judgements					

How valid are individual KPIs at judging a school's performance

Please score out of five:- 1 useless to 5 very useful

Key Stage 2 level 4 and above	
GCSE one grade G pass or more	
GCSE five grade A* - C passes or more	
GCSE point score	
A progress or value added measure	

How well do KPIs indicate performance at different levels

Please score out of five:-, 1 useless to 5 very accurate

National – How the whole Country is performing	
LEA – How individual LEAs are performing	
School – How well is a particular school performing	
Teacher – The performance of an individual teacher	

How well do KPIs indicate performance for different time scales

Please score out of five:-, 1 useless to 5 very accurate

The Past - over the last five years	
Currently - this academic year	
The Future - in five years time	

KPIs and the management process

	YES		Neither or	NO	
	Definitely	Mainly	Yes & No	Mainly	Definitely
Do KPIs help identify the key strengths of your school					
Do KPIs help identify the main weaknesses of your school					
Does your school specifically target 'borderline' pupils *(eg KS levels or GCSE C-D*					
Are KPIs used for your performance appraisal					
Do they accurately recognise the quality and value of your work					
Do KPIs play an important part in your personal and professional development					
Are you able to participate, to the degree you feel appropriate, in the setting of your targets					

k

Each pair of statements represent the extremes, please put an X in the position which best reflects your view:

How do you feel about publication of league tables

Good for education	----------	Bad for education
Increased stress	----------	Decreased stress
Accurate impression	----------	Inaccurate impression
Short-term gains	----------	Long term benefits
Fair	----------	Unfair
Ignore quality	----------	Recognise quality
Allow for local circumstances	----------	Ignore local circumstances

Other Indicators

	YES		Neither or	NO	
	Definitely	Mainly	Yes & No	Mainly	Definitely
Do schools need good numerical objective performance data					
Is it realistic to produce objective numerical data for the most important factors which indicate a school's performance					
Do you feel that you fully understand the principle of 'value added'					
Should an indication of school progress *or* value added be given in the league tables					
Should public indicators aim to compare schools with similar circumstances					
Should the relative performance of different departments within the school be made public					

m

How well do different systems judge your school's performance

n

Please score out of five:-, 1 useless to 5 very accurate

KPIs – Your current data	
Ofsted – Your last inspection report	
PANDA – The last set of data you have received	
Benchmark data – From your LEA	
Internal / Informal – Your own systems and judgements	
Other Indicator system - *Please name*	
.....	

Please make any further comments about KPIs and other Indicators, or issues raised by this questionnaire:

Thank you for taking the time and trouble to complete this questionnaire, your help with this project is very much appreciated.

Follow up: Please add your name and phone number below, if you would be willing to discuss some aspects of the research:

Name: ..... Phone:.....

Research Summary: If you would like to receive a copy of a summary please complete:

Name: ..... Email:.....

Address: .....

.....

Please return this questionnaire in the Freepost envelope:

Andy Wiggins
KPI Project
University of Durham
FREEPOST NEA242
DURHAM
DH1 1BR

Key Performance Indicators Survey

Andy Wiggins – Durham University

First of all – thank you very much for taking the time and trouble to complete the questionnaire. The number of returns was quite good for this kind of research, about 30% of schools replied, (which was similar to England). I was also very grateful for the additional comments made on many of the questionnaires. These do provide useful and valuable insight in to many of the issues, including those not referred to in the questionnaire. In addition, thanks to those who have agreed to be contacted at a later stage.

Introduction

The main purpose of this research is to assess the effect of performance indicators on school, in particular those indicators which are made public. In Scotland and England, similar secondary school data is published (eg Standard grades and GCSE). However, the situation for primary schools is very different; English results (Key stage 2 tests for 11 yr olds) are published in performance and league tables, whilst in Scotland the equivalent 5-14 tests are not. The greatest differences in the results were between English and Scottish primary schools.

There was a good distribution in the responses to how people assessed the relative position of their school compared to others in the authority, which does suggest that the returns were a reasonably representative sample. Furthermore, most respondents felt that this was a fair reflection of their overall performance. Not surprisingly there was some correlation (0.3) between these two questions:- 'better' performing schools were more likely to feel that their relative position was fairer than 'lower' performing schools.

I had expected to find some significant difference in the responses between Teachers and Heads, largely because of the changes in the role of Heads; however (and very pleasingly) this was not the case, because of this, reference is made below to schools rather than Heads or Teachers.

Details of some of the responses

Overall both primary and secondary schools felt that statutory testing systems (5-14 / Standard grades etc) are useful and will help raise standards. However, there was some evidence that these systems do encourage a 'blame culture' and worsen relationships between staff.

The results showed that parents in general made very little reference to the statutory testing system, this supports other research which suggests that parents are generally satisfied with 'their' schools. This would seem to question one of the main reasons for publishing performance data, ie that pressure from parents will drive up performance.

Not surprisingly, all schools said they were under pressure to meet their targets. Interestingly, and perhaps significantly, Scottish primary schools appeared to be under greater pressure than their English counterparts. Given that English results are made very public (newspaper league tables etc), this was surprising. It had been expected that English schools would be under more pressure. This may be partly explained by the differing roles of Education Authorities in Scotland and England. Again this brings into question the value of publishing results, even though it has been argued by various governments that doing so would increase pressure, and 'drive up' standards.

All schools were critical of the media in terms of their lack of acknowledgement of the limitations of performance data, and all Scottish primary schools were against their results being made public. It would perhaps have been reasonable to expect that 'high performing' primaries which had the most to gain, would want their results published, however this was emphatically not the case.

In both England and Scotland there was considerable evidence that the statutory indicator systems had the effect of narrowing the curriculum, ie school's tended to concentrate their efforts on what is measured. In terms of borderline pupils (those close to achieving statutory targets) there was strong evidence that schools did target these pupils[1].

In general all schools in Scotland and England did value numerical performance data, (secondary more than primary), and possibly connected to this, there was a feeling that the statutory testing systems should consider more aspects of the school.

In some of the responses there were significant differences between Scotland and England, this was most obvious at the primary level. To help quantify this, a number of the results which relate to characteristics which can be considered dysfunctional, were combined to form a 'dysfunction index', with a scale of 1 (most dysfunctional) to 5. English primary schools scored the lowest with a mean of 2.17, Scottish primaries were the highest (ie least dysfunctional effects) with a mean of 3.19[2]. Secondary schools from England and Scotland had similar results (2.72 and 2.80 respectively).

Overall the main conclusion, based chiefly on the comparison between Scottish and English primary schools, is that it would appear that publishing performance data can have a negative or dysfunctional effect on schools. Furthermore, publishing data does not appear to be a 'key driver' in terms of improving performance. Overall, these findings demonstrate the great importance of the design and management of information systems, in particular their potential dysfunctional effects.

The next step

As mentioned in the questionnaire these results will be used as part of my thesis. In addition, it is my intention to present the main findings in a paper at the European Educational Research Association conference in Edinburgh (September 2000). I will also be contacting some of the people who agreed to this nearer the summer.

Once again many thanks for your help. If you would like to make any further comments at any time, please do so:

Andy Wiggins
KPI Project
University of Durham
NEA242
DURHAM
DH1 1BR
andy.wiggins@durham.ac.uk

Please note: *This report is only intended for respondents to this survey and other interested parties. The views expressed here are based on an initial 'first look' at the data, which has not been fully analysed. None of the results or statements contained in the document should be quoted or made public without prior consent.*

---

[1] During the research period the English government introduced the 'Booster' scheme, which specifically targeted these borderline children at KS2 (end of primary).
[2] The results were statistically significant, >.001, and the effect size was >1.00.

1. Do schools need good numerical objective performance data?	YES (1) to NO (5)
2. Are the KS2 / 5 – 14 results a good judge your school's performance?	YES (1) to NO (5)
3. Do parents frequently refer to KS2 / 5 – 14 results?	YES (1) to NO (5)
4. Is there great pressure on your school to meet KS2 / 5-14 Targets?	YES (1) to NO (5)
5. Should you participate in the setting of your targets?	YES (1) to NO (5)
6. Does KS2 / 5 – 14 testing have higher credibility in your school than professional judgements?	YES (1) to NO (5)
7. Does your school concentrate on targets at the expense of other important objectives?	YES (1) to NO (5)
8. Does your school specifically target 'borderline' (KS2 / 5-14) pupils?	YES (1) to NO (5)
9. Is there any conflict between your school aims and targets?	YES (1) to NO (5)
10. How do you feel overall about KS2 / 5 – 14 testing?	Harmful (1) *to* Beneficial (5)
11. How does the target setting process affect any 'blame culture' in your school?	Encourage (1) *to* Discourage (5)
12. What do you feel the effect of KS2 / 5-14 testing is on the curriculum?	Narrowing (1) *to* Broadening (5)
13. Do you find the use of targets?	Useful (1) *to* Hindrance (5)
14. How do you feel about KS2 / 5 – 14 testing is respect of your work?	Undermining (1) *to* Supportive (5)

References

Ackoff, R. (1981) Creating The Corporate Future. John Wiley & Sons. New York.

Adab, P., Rouse, A., Mohammed, M and Marshall, T. (2002) Performance league tables: the NHS deserves better. British Medical Journal. Vol. 324:95-8.

Adams, F. (1999) 5-14: Origins, Development and Implementation. In Bryce, T and Humes, W (Eds) Scottish Education, Edinburgh University Press. Edinburgh.

Adnett, N and Davies, P. (1999) Schooling quasi-markets: Reconciling economic and sociological analyses. British Journal of Educational Studies. Vol. 47, No. 3, pp. 221-234.

Adnett, N and Davies, P. (2001) Schooling Reforms in England: From Quasi-Markets to Competition? Paper - European Education Research Association Conference. Lille.

AERA. (2000) AREA Position Statement Concerning High-Stakes Testing in Pre K-12 Education. American Educational Research Association. www.aera.net/about/policy/stakes.htm.

Ahmed, K. (2001) Blair: massive reform will deliver. The Observer. 30 December.

Aikenhead, D. (1998) Kindly businessmen, helping out schools for free. In your dreams. The Guardian. 11 September.

Aitkin, H. (1960) Taylorism at Watertown Arsenal. Harvard University Press. Boston.

Alleyne, R. (2002) Trains worse than before Hatfield. The Telegraph. 7 January.

Ansoff, H. (1965) Corporate strategy : an analytic approach to business policy for growth and expansion. McGraw-Hill. London.

Anthony, R. (1988) The Management Control Function. Harvard Business School Press. Boston.

Argyris, C. (1952) The Impact of Budgets on People. The Controllership Foundation. New York.

Argyris, C. (1960) Understanding Organizational Behaviour. Tavistock. London.

Argyris, C and Schon, D. (1978) Organizational Learning. Addison-Wesley. London.

Ashby, J and Sainsbury, M. (2001) How Do Schools Use National Curriculum Test Results. National Foundation for Educational Research. Slough.

Audit Office. (2001) FABRIC - A Framework for Performance Information. Audit Commission, Cabinet Office, National Statistic, National Audit Office and HM Treasury. March 2001.

Audit Unit. (1998) Examination Results in Scottish Schools 1996 - 98. HM Inspectors of School. The Scottish Office. Edinburgh.

Audit Unit. (1999a) Attendance and Absence in Scottish School 1995/96 to 1997/98. HM Inspectors of School. The Scottish Office. Edinburgh.

Audit Unit. (1999b) Taking a closer look at quality in schools: Links between 'How good is our school?' and the EFQM Excellence model. HM Inspectors of Schools. The Scottish Office. Edinburgh.

Baker, M. (1994) Who rules our schools. Hodder and Stougton. London.

Baker, M. (1999) Back to the future over in Scotland. Times Educational Supplement. 5 February.

Ball, S. (1990) Politics and Policy Making in Education. Routledge. London.

Ball, S. (1998) Big Policies/Small World: an introduction to international perspectives in educational policy. Comparative Education. Vol. 34, No. 2, pp. 119-130.

Barber, M . (1996) How to do the impossible: a guide for politicians with a passion for education.. Institute of Education.. London.

Barber, M. (1997) Reengineering the Political / Educational System. School Leadership and Management. Vol. 17, No. 2. pp. 187-199.

Barber, M. (1999) The Pursuit of Excellence. The Sunday Times. 7 November.

Barber, M. (2000a) Large-Scale Reform is Possible. Education Week. 15 November.

Barber, M. (2000b) High Expectations and Standards For All. No matter What: Creating a World Class Education System in England (An overview of the Blair government's approach to education reform). Paper - Smith Richardson Foundation. Washington DC.

Barber, M., Stoll, L., Mortimore, P and Hillman, J. (1995) Governing Bodies and Effective Schools. DFEE. London.

Barnard, N. (1999) Results rise steady as Blair cools on tables. Times Educational Supplement. 26 November.

Bartholomew, D. (1999) Process is Back. Industry Week. 1 November.

Bassey, M. (2001) Delusions in the politics of education. Research Intelligence. No. 75.

Batteson, C. (1997) A review of politics of education in the 'Moment of 1976'. British Journal of Educational Studies. Vol. 45, No. 4, December, pp 363-377.

Bazalgette, J. (1996) Greater then a mere sum of skills. Times Educational Supplement. 31st May.

BBC. (1999) The train not stopping at platform one.... BBC. 24 March.

BBC. (2001) League Tables Ups and Downs. BBC. 5 December.

BBC. (2002a) Education chief attacks test regime. BBC. 11 April.

BBC. (2002b) Blair takes on public service 'wreckers'. BBC. 3 February.

BBC Scotland. (2001) Concern over school league tables. BBC. 29 November.

Beare, H., Caldwell, B and Millkan, R. (1989) Creating an excellent school. Routledge. London.

Bell, L. (1998) From Symphony to Jazz: the concept of strategy in education. School Leadership and Management. Vol. 18, No. 8, pp. 449-460.

Bennett, N., Crawford, M., Levacic, R., Glover, D and Early, P. (2000) The Reality of School Development Planning in the Effective Primary School: technicist or guiding plan?. School Leadership and Management. Vol. 20, No. 3, pp. 333-351.

Birnberg, J., Turopolec, L and Young, M. (1983) The Organizational Context of Accounting. Accounting Organizations and Society. Vol. 8, No. 2/3, pp. 111-129.

Blair, T. (1998) Speech given at Ruskin College, Oxford. www.leeds.ac.uk/educol/documents/000000084.htm. 27 June.

Blair, T. (2001) Address to conference for new headteacher. National College for School Leadership. Video presentation, www.ncsl.org.uk, 13 November.

Blair. T and Blunkett. D. (2000) National College for School Leadership - A Prospectus. DfEE. London .

Bloomer, K. (1999) The Local Governance of Education: An Operational Perspective. In Bryce, T and Humes, W (Eds) Scottish Education, Edinburgh University Press. Edinburgh.

Blunkett, D. (1999) Human Capital The Key To Business Future (Press release 115/99). DfEE . 11 March.

Blunkett, D. (2000) Achieving Excellence Through The National Curriculum. Letter to QCA. www.qca.org.uk/ncr/db-letter.htm accessed 10/12/00.

Bogdanor, V. (1979) Power and Participation. Oxford Review of Education. Vol. 5. pp.157-68.

Boland, T and Fowler, A. (2000) A systems perspective of performance management in public sector organisations. International Journal of Public Sector Management. Vol. 13, No. 5. pp. 417-446.

Bostingl, J. (1993) The quality movement: what's it really about. Educational Leadership. No. 51. pp. 66.

Bottery, M. (1994) Education and the Convergence of Management Codes. Educational Studies. Vol. 20, No. 3, pp. 329-345.

Bottery, M. (1999) Global Forces, National Mediations and the Management of Educational Institutions. Educational Management and Administration. Vol. 27, No. 27, pp. 299-312.

Briers, M and Hirst, M. (1990) The role of budgetary information in performance evaluation. Accounting Organizations and Society. Vol. 15, No. 4, pp. 373-98.

Brignall, S and Modell, S. (2000) An institutional perspective on performance measurement and management in the 'new' public sector. Management Accounting Research. Vol. 11, pp. 281-306.

British Medical Association. (1999) BMA calls for waiting times to be key indicator. BMA Press release. 15 February.

British Medical Association. (2000) Clinical Indicators: (League Tables). BMA. March.

Broadbent, J., Dietrich, M. and Laughlin, R. (1996) The Development of Principal-Agent, Contracting and Accountability Relationships in the Public Sector: Conceptual and Cultural Problems. Critical Perspectives on Accounting. Vol 7, pp. 259-284.

Broadbent, J., Jacobs, K and Laughlin, R. (1999) Comparing schools in the UK and New Zealand: individualizing and socializing accountabilities and some implications for management control. Management Accounting Research. Vol. 10, pp. 339-361.

Broadfoot, P. (1999) Empowerment or Performativity? English Assessment Policy in the late twentieth Century. Paper - British Educational Research Association Conference. Brighton, September.

Brundrett, M. (1999) The National Professional Qualification for Headship: perceptions of the providers of taught higher degrees in educational management in England and Wales. School Leadership and Management. Vol. 19, No. 4, pp. 497-510.

Brundrett, M. (2000) The Question of Competence: the origins, strengths and inadequacies of a leadership training paradigm. School Leadership and Management. Vol. 20, No. 3, pp. 353-369.

Bryce, T and Humes, W. (1999) Scottish Education. Edinburgh University Press. Edinburgh.

Buckley, A and McKenna, E. (1972) Budgetary control and business behaviour. Accounting and Business Research. Spring, pp. 137-150.

Budge, D. (2000) Private troubles nag model academies. Times Educational Supplement. 16 June.

Burns, J, Joseph,N, Lewis, L Scapens, R, Southworth, A and Turley, S. (1994) Accounting Standards and Management Decisions: A survey of UK Management Accountants. European Accounting Association Annual Congress April 1994. .

Burns, W.J and De Coster, D.T. (1969) Accounting and its Behavioural Implications. McGraw-Hill. New York.

Bush, T. (1989) Managing Education: Theory and Practice. Open University Press. Milton Keynes.

Bush, T. (1998) The National Professional Qualification for Headship: the key to effective school leadership?. School Leadership and Management. Vol. 18, No. 3. pp. 321-333.

Bush, T. (1999) Crisis or Crossroads? The Discipline of Educational Management in the Late 1990s. Educational Management and Administration. Vol. 27, No. 3. pp. 239-252.

Bush, T and Coleman, M. (2000) Leadership and Strategic Management in Education. Educational Management Development Unit, University of Leicester. Paul Chapman. London.

Cabinet Office. (1991) The Citizen's Charter: Raising the standard (Cm1599). Cabinet Office. London.

Cabinet Office. (1995) Parent's Charter - Scotland. Cabinet Office. London.

Cabinet Office. (1999a) Performance Management: Civil Service Reform - A Report to the Meeting of Permanent Heads of Departments, Sunningdale. Cabinet Office. London.

Cabinet Office. (1999b) Modernising government - White paper (Cm 4310). Cabinet Office. London.

Cabinet Office. (2000a) Public Sector Benchmarking Project - Introduction. Cabinet Office. London.

Cabinet Office. (2000b) Minister praises centres of excellence in Oldham and Cumbernauld. Press release CAB 319/00, 4 October. London.

Callaghan, J. (2001) Towards and national debate (Re-print of the full text) . The Guardian. 15 October.

Capper, C and Jamison, M. (1993) Let the buyer beware: total quality management and educational research and practice. Educational Researcher. November, pp. 25-30.

Carr, D and Johnson, H. (1995) Best Practice Reengineering. McGraw-Hill. New York.

Carson, P., Lanier, P., Carson, K and Birkenmeier, B. (1999) A historical perspective on fad adoption and abandonment. Journal of Management History. Vol. 5. No. 6. pp.320-333.

Carter, N, Klein, R and Day, P. (1992) How organisations measure success : the use of performance indicators in government . Routledge. London.

Carvel, J. (1998) Blunkett Targets Poor School Zones. Guardian. 5 December.

Carvel, J and Smithers, R. (1998) Tables warp teaching, warn heads. Times Educational Supplement . 1 December.

Carver, G. (2000) Tested to Destruction?: a survey of examination stress in teenagers. Professional Association of Teachers. London.

Cassidy, S. (1999) Official leaked English Tests. Times Educational Supplement. 2 July.

Caulkin, S. (1996) Business fashion: what's in and out. The Observer. 19 May pp6.

Caulkin, S. (2000) Thames Valley or Rhine Valley. The Observer. 7 May.

Cave, E. (1990) Local Management of Schools: Some Practical Issues. Routledge. London.

Chalmers, I. (1998) Unbiased, relevant, and reliable assessments in health care. British Medical Journal. Vol. 317: 1167-1168.

Chandler, J. (1996) The Citizen's Charter. Dartmouth. Aldershot.

Chappell, R. (1994) Can TQM in Public Education Survive Without Co-Production?. Quality Progress. Vol. 27, No. 7, pp. 41-44..

Charter, D and Wilkinson, P. (1996) Sacking to be investigated. The Times. 24 October.

Chartered Institute of Public Finance and Accountancy. (1988) Performance Indicators in Schools. CIPFA. London.

Chartered Institute of Public Finance and Accountancy. (1998) Accountability - A Framework for Public Services. CIPFA. London.

Cheek, D., Fitz-Gibbon, C and Tymms, P. (2000) The English Reforms Are Not For Us. Education Week. 13 December.

Chenhall, R, Harrison G and Watson, D (Eds). (1981) The Organisational Context of Management Accounting. Pitman. Melbourne.

Child, J. (1969) British management thought. Allen and Unwin. London.

Children's mercy hospital. (2002) Physician Performance Data - Control Charts. www.childrens-mercy.org/stats.

Chitty, C. (1989) Towards a New Education System: The Victory of the New Right. The Falmer Press. Lewes.

Chitty, C. (1998) The 'Moment of 1976' Revisited. British Journal of Educational Studies. Vol. 46, No. 3, September, pp 318-323.

Chrisafis, A. (2002) Widespread cheating devalues school tests. The Guardian. 28 October.

Coe, R. (1999) Changes in Examination Grades over Time: Is the same worth less? Paper - British Educational Research Association, Conference. Brighton, September.

Coe, R. (2001) How safe is that PANDA? A comparison between DfEE benchmarking and other ways of contextualising school performance. Paper - British Educational Research Association, Conference. Leeds, September.

Cohen, J. (1977) Statistical Power Analysis for the Behavioural Sciences. Academic Press. New York.

Cohen, M. (1981) Effective schools: What the research says. Today's Education, April/May. 46G-49G.

Collier, P. (2001) Police Performance Measurement and Human Rights. Public Money & Management. Vol. 21, No. 3, pp. 35-39.

Cook, T and Campbell, D. (1979) Quasi-Experimentation: Design and Analysis Issues for Field Settings. Rand McNally. Chicago.

Cope, S., Leishman, F., Starie, P. (1997) Globalization, new public management and the enabling State. International Journal of Public Sector Management. Vol. 10 No. 6 , pp 444-460.

Cowie, M. (2000) Talking Heads: A Critical Analysis of the Quality Assurance Relationship between Secondary Schools and an Education Authority. Paper - European Educational Research Conference. Edinburgh, September.

Cramer, D. (1994) Introducing Statistic for Social Research. Routledge. London.

Crawford, L and Shutler, P. (1999) Total quality management in education: problems and issues for the classroom teacher. The International Journal of Educational Management. Vol. 13, No. 3, pp. 67-72.

Crooks, T., Kane, M and Cohen, A. (1996) Threats to the valid use of assessments. Assessment in Education. Vol. 3, No. 3, pp. 265-285.

Croxford, L. (1999) League Tables - Who Needs Them?. CES Briefings. Edinburgh.

Croxford, L. (2000) Is Free-Meal Entitlement a Valid Measure of School Intake Characteristics?. Educational Research and Evaluation. Vol. 6, No. 4, pp. 317-335.

Croxford, L and Cowie, M. (1996) The Effectiveness of Grampian Secondary Schools. Centre for Educational Sociology. Edinburgh.

Cutler, T and Waine, B. (1999) Rewarding Better Teachers?. Educational Management and Administration. Vol. 27, No. 1, pp. 55-70.

Daugherty, R. (1995) National Curriculum Assessment: A Review of Policy. Falmer. London.

Davies, B. (1997) Reengineering and its Application to Education. School Leadership and Management. Vol. 17, No. 2, pp. 173 - 185.

Davies, B and Ellison, L. (1998) Strategic Planning in Schools: an oxymoron? School Leadership and Management. Vol. 18, No. 4, pp. 461-473.

Davies, B and Ellison, L. (1999) Strategic Direction and Development of the School. Routledge. London.

Davies, H and Lampel, J. (1998) Trust in performance indicators?. Quality in Health care. 7:159-162.

Davies, H., Nutley, S and Smith, P. (1999) Editorial: What works? The Role of Evidence in Public Sector Policy and Practice. Public Money and Management. January-March, pp. 3-5.

Davies, P. (1999) What Is Evidence-Based Education? British Journal of Educational Studies. Vol. 47, No. 2, pp. 108-121.

Davis, A. (1999) Educational Assessment: a critique of current policy. Philosophy of Education Society of Great Britain. Keele.

De Cock, C and Hipkin, I. (1997) TQM and BPR: Beyond the Beyond Myth. Journal of Management Studies. Vol. 34, No. 5, pp. 659-675.

Dearden, J. (1987) Measuring Profit Centre Managers. Harvard Business Review. Sept - Oct .

Dearlove, D. (1997) Officers train to be hot shots in personnel. The Times. 30 January.

Demaine, J. (1999) Education Policy and Contemporary Politics. Macmillan. Basingstoke.

Deming Organisation. (2000) Dr. W. Edwards Deming . http://deming.org/deminghtm/deming.html.

Deming, W. (1986) Out of the Crisis. Cambridge University Press. Cambridge.

Dent, H. (1977) Education in England and Wales. Hodder and Stougton. London.

Dent, J. (1996) Global competition: challenges for management accounting and control. Management Accounting Research. Vol. 7, pp. 247-269.

Department for Education. (1994) Our Children's Education: The Updated Parents Charter. HMSO. London.

Department for Education and Employment. (1997) Excellence in Schools (Cm 3681). HMSO. London.

Department for Education and Employment. (1998a) Learning and working together for the future: a strategic framework to 2002. DfEE. London.

Department for Education and Employment. (1998b) Target-setting in schools. DfEE. Circular 11/98. London.

Department for Education and Employment . (1999) National College for School Leadership - A Prospectus. DFEE. London.

Department for Education and Employment. (2000a) Recognising Progress – Getting the most from your data (Ref 0253). DfEE. London.

Department for Education and Employment. (2000b) The Autumn Package 2000: Pupil Performance Information (0257). DfEE. London.

Department for Education and Skills. (2001e) School Achievement Awards Scheme – Round 2 (0675). DfES. London.

Department for Education and Skills. (2001a) Schools achieving success. DfES. London.

Department for Education and Skills. (2001b) Final decision on changes to the performance tables in 2001 (0578). DfES. London.

Department for Education and Skills. (2001c) Guidance on the Threshold Process (0538). DfES . London.

Department for Education and Skills. (2001d) Delivering Results: A strategy to 2006. DfES. London.

Department for Education and Skills. (2001f) Performance Management. DfES. London.

Department for Education and Skills. (2001g) Morris wants business backing to bridge skills gap (0372). DfES. London.

Department for Education and Skills. (2002) Final decision on changes to the performance tables in 2002 (0536). DfES. London.

Department of Education and Science. (1988) Education Reform Act. HMSO. London.

Department of Education and Science. (1989) Performance Indicators: An aide-memoire from the DES. In Education. 8 December, pp. 514-515.

Department of Education and Science. (1991) Development Planning: A Practical Guide. HMSO. London.

Department of Education for Northern Ireland. (2001) School Performance Tables Not To Be Published - McGuinness. Press Release. 10 January.

Department Of Health. (2002) Performance Indicators and Clinical Indicators: Part A Technical Specifications. Department of Health. www.doh.gov.uk/nhsperformanceindicators/2002/haed_t.doc.

Devine, M. (1994) Riding the re-engineering tiger. The Sunday Times. 26 June.

Diggins, P. (1997) Reflections on Leadership Characteristics Necessary to Develop and Sustain Learning School Communities. School Leadership and Management. Vol.17, No.3, pp. 413-425.

Dobson, R. (1999) Survival alone is not adequate measure of success, Bristol inquiry told. British Medical Journal. Vol. 318:832; 27 March.

Docking, J. (1996) National School Policy. David Fulton. London.

Doe, B. (1999) The Role of the Media in the Reform of School Management. Educational Management and Administration. Vol. 27, No. 3, pp. 335-343.

Doherty, N and Mistry, A. (1996) The application of Business Process Re-engineering in the United Kingdom: an empirical analysis. Paper - European Academic Conference on Business Process Re-engineering. Canfield.

Doll, R. (1998) Controlled trials: the 1948 watershed. British Medical Journal. Vol. 317: 1217-1200.

Dorn, S. (1998) The Political Legacy of School Accountability Systems. Education Policy Analysis Archives. Vol. 6, No. 1.

Drucker, P. (1954) The Practice of Management. Harper & Row. New York.

Drucker, P. (1973) Management. Harper & Row. New York.

Drucker, P. (1976) What Results Should You Expect? A Users' Guide to MBO. Public Administration Review. Vol. 36, (January / February), pp. 12-19.

Drucker, P. (1977) Management :Tasks, Responsibilities, Practices. Pan. London.

Drucker, P. (1990) Managing the Non-Profit Organization. Butterworth-Heinemann. Oxford.

Drucker, P. (1999) Management challenges for the 21st century. Butterworth-Heinemann. Oxford.

Drury, C and Tayles M. (1997) Evidence on the Financial Accounting Mentality Debate: A Research Note. British Accounting Review . pp. 263 - 276.

Earl, L., Levin, B., Leithwood, K., Fullan, M and Watson, N. (2001) Watching and Learning 2: OISE/UT Evaluation of the Implementation of the National Literacy and Numeracy Strategies (DfES 0617/2001). Ontario Institute for Studies in Education, University of Toronto. London.

Eccles, R. (1991) The Performance Measurement Manifesto. Harvard Business Review. Jan - Feb 1991.

Edwards, M. (1996) Client orientation in a first school. in Fidler, B (Ed) Strategic Planning for School Improvement. Pitman. London.

Edwards, P., Ezzamel, M., Robson, K. (1999) Connecting Accounting and Education in the UK. Critical Perspectives on Accounting. Vol. 10, 469-500.

Elcock, H. (1996) What Price Citizenship? Public Management and the Citizen's Charter. In Chandler, J (Ed) The Citizen's Charter. Dartmouth. Aldershot.

Elliott, L. (2001) Wales's end-of-year report card. BBC. 26 December.

Emmanuel, C, Otley, D and Merchant, K. (1995) Accounting and Management Control. Chapman and Hall. London.

Emmanuel, C, Otley, D and Merchant, K (Eds). (1992) Readings in Accounting and Management Control. Chapman and Hall. London.

Engineering Council. (2000) Measuring the Mathematics Problem. Engineering Council. London. June.

Ezzamel, M., Hoskin, K and Macve, R. (1990) Managing it all by numbers: a review of Johnson and Kaplan's 'Relevance Lost'. Accounting and Business Research. Spring, pp. 153-166.

Farrell, C. (1999) The Accountability of School Governing Bodies. Educational Management and Administration. Vol. 27, No.1, pp. 119-127.

Farrell, C and Morris, J. (1999) Professional Perceptions of Bureaucratic Change in the Public Sector: GPs, Headteachers and Social Workers. Public Money & Management. Vol. 19. No. 4, pp. 31-36.

Fidler, B. (1996) Strategic Planning for School Improvement.. Pitman. London.

Fielding, M. (1999) Taking Education Really Seriously: two years hard labour. Cambridge Journal of Education. Vol. 29, No. 2, pp. 173-181.

Fink, A. (1995) How to analyze survey data. Sage. Thousand Oaks.

Fitz, J. (1999) Reflections on the Field of Educational Management Studies. Educational Management and Administration. Vol. 27, No. 3, pp. 313-321.

Fitz-Gibbon, C. (1990) Performance Indicators: A Bera Dialogue. Avon Multi-lingual Matters. Clevedon.

Fitz-Gibbon, C. (1996) Monitoring Education. Indicators, Quality and Effectiveness. Cassell. London.

Fitz-Gibbon, C. (1997) The Value Added National Project: Final Report. School Curriculum and Assessment Authority. London.

Fitz-Gibbon, C and Vincent, L. (1997) Difficulties Regarding Subject Difficulties: developing reasonable explanations for observable data. Oxford Review of Education. Vol. 23, No. 3, pp. 291-298.

Flanagan, J. (2002) It's official: Hackney is more dangerous than Soweto. The Telegraph. 14 April.

Fletcher, C. (1993) Appraisal: an idea whose time has gone. Personnel Management. Vol. 25, No. 9, pp. 34-37.

Fowler, F. (2002) Survey Research Methods. Sage. London.

Freeman, R. (1993) Quality Assurance in Training and Education. Kogan Page. London.

Frey, J. (1989) Survey Research by Telephone. Sage. London.

Gaffney, D., Pollock, A., Price, D and Shaoul, J. (1999) The Politics of the private finance initiative and the new NHS. British Medical Journal. Vol. 319, 24 July.

Gallagher, J. (2002) What Next for OERI. Education Week. 27 March.

Gambles, I. (1999) A balanced future lies in the cards. Public Finance. 16 April.

Garner, R. (2001) Private firm fined £300,000 for failing to improve results. The Independent. 1 September.

Gibson, A and Asthenia, S. (1999) Schools, markets and equity: access to secondary education in England and Wales. Paper - AERA conference. Montreal, April.

Gipps, C. (1994) Beyond Testing: Towards a Theory of Educational Assessment. Falmer. London.

Glatter, R. (1987) Towards and Agenda for Educational Management. Educational Management and Administration. Vol. 15, No. 1, pp. 5-12.

Glatter, R. (1999) From Struggling to Juggling: Towards a Redefinition of the Field of Educational Leadership and Management. Educational Management and Administration. Vol. 27, No. 3, pp. 253-266.

Goddard, A. (1997) Organisational Culture and Budgetary Control in a UK Local Government Organisation. Accounting and Business Research. Spring.

Goddard, M and Smith, P. (2001) Performance measurement in the new NHS. Health Policy Matters, . University of York.

Gold, K. (2002) Great results fudge true picture. Times Educational Supplement. 10 May.

Goldstein, H and Myers, K. (1996) Freedom of information: towards a code of ethics for performance indicators. Research Intelligence. Vol. 57, pp. 12-16.

Goldstein, H. and Spiegelhalter, D. (1996) League tables and their limitations: statistical issues in the comparisons of institutional performance. Journal of the Royal Statistical Society. A, 159, pp 385-443.

Goldstein, H., Rash bash, J., Yang, M., Woodhouse, G., et al. (1993) A Multilevel analysis of school examination results. Oxford Review of Education. Vol. 19, pp. 425-433.

Goodbody, J. (1999) Why PE is in decline. The Times. 27 October.

Gorard, S. (2000) Questioning the crisis account: a review of evidence for increasing polarization in schools. Educational Research. Vol. 42, No. 3, pp. 309-321.

Gorard, S. (2002) Political Control: A way forward for Educational Research?. British Journal of Educational Studies. Vol. 50, No. 3, pp. 378-389.

Gorz, A. (1989) Critique of Economic Reason. Verso. London.

Grace, G. (1995) School Leadership: Beyond Education Management. The Falmer Press. London.

Grace, G. (2000a) Research and the challenges of contemporary school leadership: The contribution of critical scholarship. British Journal of Educational Studies. Vol. 48, No. 3, pp. 231-247.

Grace, J. (1999) Stressed out - at the age of seven. The Guardian. 20 July.

Grace, J. (2000b) Literacy 2, Sport 0. The Guardian. 29 February.

Granström, K. (1996) Decentralization and teachers: Professional status cannot be granted, it has to be acquired. In J. Chapman, W. Boyd, R. Lander and D. Reynolds (eds) The Reconstruction of Education: Quality, Equality and Control. Cassell. London.

Greenwood, J. and Wilson, D. (1989) Public Administration in Britain Today. Unwin Hyman. London.

Grey, A. (1997) New Public Policy Management and the Need for Evidence: the function of evaluation. Paper - Evidence-Based Policies and Indicator system, Conference. Durham.

Grint, K. (1994) Reengineering history: social resonance's and business process reengineering. Organization. Vol. 1, pp. 179-201.

Grinyer, J., Russell, A and Collinson, D. (1998) Evidence of Managerial Short-termism in the UK. British Journal of Management. Vol. 9, pp. 13-22.

Gunter, H. (1996) Appraisal and the Schools as a Learning Organisation. School Organisation. Vol. 16, No. 1, pp 89-100.

Hackett, G. (2000) Junior targets to be imposed by ministers. Times Educational Supplement . 28 July.

Hakim, C. (2000) Research Design: successful designs for social and economic research. Routledge. London.

Halachmi, A and Bovaird. (1997) Process reengineering in the public sector: learning some private sector lessons. Technovation. Vol. 17, No. 5, pp. 227-235.

Hallgarten, J. (2002) 23,000 League Tables Under the Sea? Institute for Public Policy Research. London.

Hamel, G and Prahalad, C. (1994) Competing for the future / by Gary . Harvard Business School Press. Boston.

Hammer, M and Champy, J. (1993) Reengineering the Corporation. Harper. New York.

Haney, W. (2000) The myth of the Texas miracle in education. Education Policy Analysis Archives. Vol. 8, No. 41.

Hargreaves, D. (1994) The Mosaic of Learning: schools and teachers for the new century. Demos. London.

Hargreaves, D and Hopkins, D. (1991) The Empowered School. Cassell. London.

Harlen, W and Deakin Crick, R . (2002) A systematic review of the impact of summative assessment and tests on students' motivation for learning (EPPI-Centre Review, version 1.1). In: Research Evidence in Education Library, EPPI-Centre, Social Science Research Unit, Institute of Education. London.

Harris, M., Smith, D and Champagne, D. (1995) A field study of performance appraisal purpose: research - versus administrative-based ratings. Personnel Psychology. Vol. 48, No. 1, pp. 151-161.

Harris, R. (1999) Blair's master class for Machiavelli. The Sunday Times. 17 January.

Harrison, S. (1995a) The effect of externally-imposed instruments on the performance of public sector organisations. Paper - British Academy of Management Conference. Sheffield, September.

Harrison, S. (1995b) Understanding the information needs of the local community: the local versus the central government perspective. Paper - CERP-Education Research Committee. Maastricht.

Harrison, S. (1999) Propaganda, Persuasion and Symmetry: Local and Central Government Perspectives on Communicating with the Citizen. British Journal of Management. Vol. 10, pp.53-62.

Hartley, D. (1990) Tests, tasks and Taylorism: a model T approach to the management of education. Journal of Education Policy. Vol. 5, No. 1, pp. 67-67.

Hartley-Brewer, E. (2001) Learning to Trust and Trusting to Learn: How schools can affect children's mental health. Institute of Public Policy Research. July.

Harvey, D. (1995) Re-engineering: the critical success factors. Management Today. London.

Hattersley, R. (1998) Apartheid grammar poll hits parents. Times Educational Supplement . 20 November.

Hay McBer. (2000) The Lessons of Leadership. Hay McBer. December.

Helsby, G and Saunders, M. (1993) Taylorism, Tylerism and Performance Indicators: defending the indefensible?. Educational Studies. Vol. 19, No. 1, pp. 55-77.

Henderson, D. (2001) No league tables - McConnell. Scottish Times Educational Supplement. 23 February.

Henderson, D. (2002) Ministers turn away from baseline tests. Times Educational Supplement. 22 March.

Henry, J. (2001) Heads mistrust new 'value-added' scores. Times Educational Supplement. 23 November.

Hequet, M. (1994) Giving good feedback. Training. Vol. 31, No. 9, pp. 72-78.

Her Majesty's Inspectorate of Constabulary. (2000) On the Record: Thematic Inspection Report on Police Crime Recording. Home Office. London.

Heyward, J., Siebert, W., Xiangdong, W. (1997) Payment by Results Systems: British Evidence. British Journal of Industrial Relations. 35:1 March 0007-1080 pp. 1-22.

Hillage, J., Pearson, R., Anderson, A and Tamkin, P. (1998) Excellence in Research in Schools. DfEE. London.

Hofstede, G. (1968) The Game of Budget Control. Koninklijke Van Gorcum. The Netherlands.

Holmes, G and McElwee, G. (1995) Total quality management in higher education: how to approach human resource management. The TQM Magazine. Vol. 7, No. 6, pp. 5-10.

Home Office. (2001) National Crime Figures. Research, Development and Statistics Directorate. 19 July

Hood, C. (1991) A public management for all seasons?. Public Administration. Vol. 69 No. 1.

Hope, J and Fraser, R. (1997) Beyond budgeting...Breaking through the barrier to the 'third way'. Management Accounting. December, pp. 20-23.

Hopper, T., Kirkham, L., Scapens, R and Turley, S. (1992) Does financial accounting dominate management accounting - a research note. Management Accounting Research. Vol. 3, No. 4, pp. 307-311.

Hopwood, A. (1972) An Empirical Study of the Role of Accounting Data in Performance Evaluation. Journal of Accounting Research Supplement. pp. 156-163.

Hopwood, A. (1976) Accounting and Human Behaviour. Prentice-Hall. New Jersey.

Hopwood, A.G and Miller, P (Eds). (1994) Accounting as Social and Institutional Practice. Cambridge University Press. Cambridge.

House of Commons. (2000) School League Tables Debate. Hansard, 26 January, Column 47WH to 69WH. London.

Howard, M. (2002) Labour way off on public service reform. www.conservatives.com news. 18 November.

Hoyle, B. (2002) New figures reveal shocking increase in rural crime. The Telegraph. 21 April.

Huczynski, A., and Buchanan, D. (1991) Organizational Behaviour. Prentice Hall. Hemel Hempstead.

Hughes, M. (1986) Theory and Practice in educational management. In Hughes, M., Ribbins, P and Thomas, H. (Eds), Managing Education: The System and the Institution. Holt, Rinehart and Winston. London.

Hughes, O. (1994) Public Management and Administration. St Martin's Press INC. New York.

Jack, I. (2001) The crash that stopped Britain. Granta. London.

Jackson, P. (1988) The management of performance in the public sector. Public Money & Management. Vol. 8, pp. 11-16.

Jacobs, K and Manzi, T. (2000) Performance indicators and social constructivism. Critical Social Policy. Vol. 20, No. 1, pp. 85-103.

James, M. (2000) Measured lives: the rise of assessment as the engine of change in English schools. The Curriculum Journal. Vol. 11, No. 3, pp. 343-364.

Jaques, E. (1990) In Praise of Hierarchy. Harvard Business Review. January-February 1990.

Johnson, G and Scholes, K. (1993) Exploring Corporate Strategy. Prentice Hall. Hemel Hempstead.

Johnson, H. (1994) Relevance regained: total quality management and the role of management accounting. Critical Perspectives on Accounting. Vol. 5. No. 3, pp. 259-267.

Johnson, H. (1996) Performance Measurement and Targets. in J Docking (ed) (1996) National School Policy: Major Issues in Education Policy for Schools in England and Wales. 1979 onwards. David Fulton, London.

Johnson, H and Kaplan, R. (1991) Relevance Lost: The Rise and Fall of Management Accounting. Harvard Business School Press. Boston.

Johnson, J. (2001) Cheating more widespread in primary schools. The Guardian. 3 September.

Jones, C and Dewing, I. (1996) Management Accounting Information and Organisational Change in a NHS Trust Hospital. CIMA. London.

Jones, R and Keasey, K. (2000) Financial Services League Tables. University of Leeds. Working paper.

Jones, T and Dugdale, D. (1995) Manufacturing accountability, in Berry, A. Broadbent, J. Otley, D. (eds) Management Control. Macmillan. London.

Jowett, P. and Rothwell, M. (1988) Performance Indicators in the Public Sector. Macmillan. Basingstoke.

Kaplan, R. and Norton, D. (1996) The Balanced Scorecard. translating strategy into action.. Harvard Business School Press. Boston.

Keasey, K., Moon, P and Duxbury, D. (2000) Performance measurement and the use of league tables: some experimental evidence of dysfunctional consequences. Accounting and Business Research. Vol. 30, No. 4, pp. 275-286.

Kelly, J. (1982) Scientific Management: Job Redesign and Work Performance. Academic Press. London.

Kerlinger, F. (1986) Foundations of behavioural research. CBS Publishing. New York.

Kingdom, J. (1996) Citizen or State Consumer? A Fistful of Charters. In Chandler, J (Ed) The Citizen's Charter, Dartmouth. Aldershot.

Kirkland-Rowell. (2000) Education perception monitor. www.kirkland-rowell.com. .

Klein, S., Hamilton, L., McCaffrey, D and Stecher. B. (2000) What Do Test Scores in Texas Tell Us?. Education Policy Analysis Archives. Vol. 8, No. 49.

Koontz, H. (1961) The Management Theory Jungle. Academy of Management Journal. No. 4, pp. 174-188.

Koontz, H . (1980) The Management Theory Jungle Revisited. Academy of Management Review. No. 5, pp. 175-187.

Koontz, H., O'Donnell, C., and Weihrich, H. (1980) Management. McGraw-Hill. Tokyo.

Kotler, P. and Armstrong, G. (1992) Principles of Marketing. Prentice Hall International. London.

Kouzim, A. Loffler, E., Klages, H and Korac-Kakabadse, A. (1999) Benchmarking and performance measurement in the public sectors. The International Journal of Public Sector Management. Vol. 12, No. 2, pp. 121-144.

Lan, Z and Rosenbloom, D. (1992) Editorial. Public Administration Review. Vol. 52, pp. 535-7.

Laughlin, R., Broadbent, J., Shearn, D. and Willig-Atherton, H. (1994) Absorbing LMS: The Coping Mechanism of a Small Group. Accounting, Auditing & Accountability Journal. Vol. 7, No. 1, pp.59-85.

Lawrence, P and Lorsch, J. (1967) Organizations and Environment. Harvard Press. Boston.

Lawton, A and Rose, A. (1991) Organisation and Management in the Public Sector. Pitman. London.

Leach, R. (1996) The Charter and Education. In Chandler, J (Ed) The Citizen's Charter, Dartmouth. Aldershot.

Leicester, G. (1999) The Seven Enemies of Evidence-Based Policy. Public Money & Management. January - March, pp. 5-7.

Levacic, R and Glatter, R. (2001) Really Good Ideas? Developing Evidence-Informed Policy and Practice in Educational Leadership and Management. Educational Management and Administration. Vol. 29, No. 1, pp. 5-25.

Lumpkin, G and Dress, G. (1995) Simplicity as a Strategy-Making Process: The Effects of Stage of Organizational Development and Environment on Performance. Academy of Management Journal. Vol. 38, No. 5, pp. 1386.

Lynch, R. and Cross, K. (1995) Measure Up: How to Measure Corporate Performance. Blackwell. Oxford.

Macintosh, N. (1985) The Social Software of Accounting and Information Systems. John Wiley & Sons. New York.

Mahony, P and Moos, L. (1998) Democracy and School leadership in England and Denmark. British Journal of Educational Studies. Vol. 46, No. 3, pp. 302-317.

Mansell, W. (2000) GTC's 52 varieties of professional conduct. Times Educational Supplement . 15 September.

Marcus, J. (2000) What the doctor ordered. Times Educational Supplement. 24 June.

Marsden, D and Richardson, R . (1994) Performing for pay? The effect of 'merit pay' on motivation in a public service. British Journal of Industrial Relations. Vol.32, No.2, pp. 243-262.

Marshall, E and Spiegelhalter, D. (1998) Reliability of league tables in vitro fertilisation clinics: retrospective analysis of live birth rates. British Medical Journal. Vol. 316:1701-1705.

Marsushita, K. (1982) Quoted in Durham MBA 1995 – Strategy Unit. Strategy module. .

Mayo, J. (2000) Effective reporting in education: Enabling local management and focussing

investment to achieve national targets. HM Treasury: Public Services Productivity Panel. 27 July.

McBain. B and Louden, D. (1998) Hard work pays off. The Herald. 26 November.

McCartney, S. and Brown, R. (1999) Managing by numbers: using outcome measures in the NHS. International Journal of Health Care Quality Assurance. Vol. 12. No. 1. pp 6-12.

McGregor, D. (1957a) The Human Side of Enterprise. Management Review. November 1957,. in Shafritz, J and Ott, J. (1992) Classics of Organization Theory. Wadsworth. Belmont.

McGregor, D. (1957b) An Uneasy Look at Performance Appraisal. Harvard Business Review. Vol. 35, No. 3, pp. 89-94.

McGregor, D. (1960) The Human Side of Enterprise. McGraw-Hill. New York.

Meredith, P. (2000) The contracting out of LEA functions and implications for democracy and accountability. Education and the Law. Vol. 12, No. 1, pp. 5-7.

Merson, M. (2000) Teachers and the myth of modernisation. British Journal of Educational Studies. Vol. 48, No. 2, pp. 155-169.

Meyer, B. (1995) Natural and Quasi-Natural Experiments in Economics. Journal of Business and Economic Statistics. Vol. 13, No. 2, pp. 151-161.

Miles, A and Huberman, M. (1984) Qualitative data analysis : a sourcebook of new methods. Sage. London.

Miller, D. (1991) Handbook of Research Design and Social Measurement. Sage. Newbury Park.

Mintzberg, H. (1993) Structures in Fives: Designing Effective Organizations. Prentice Hall. New Jersey.

Mintzberg, H. (1994) The Rise and Fall of Strategic Planning. The Free Press. New York.

Mirron, G and Applegate, B . (2000) An Evaluation of Student Achievement in Edison Schools. The Evaluation Center, Western Michigan University,. www.wmich.edu/evalctr.

Monbiot, G. (2000) The Captive State: The Corporate Takeover of Britain. Macmillan. London.

Monbiot, G. (2002) Public Fraud Initiative. The Guardian. 18 June.

Moon, P and Fitzgerald, L. (1996) Delivering the good at TNT: the role of performance measurement system. Management Accounting Research. Vol. 7. pp. 431-457.

Mooney, T. (1999) The 11+ stalks our schools: Are institutions cheating to ensure a brighter intake?. The Guardian. 6 July.

Morgan, P and Allington, N. (2001) Has the Public Sector Retained its 'Model Employer' Status?. Public Money & Management. Vol. 22. No. 1. pp. 35-42.

Morley, L and Rassool, N. (2000) School effectiveness: new managerialism, quality and the Japanization of education. Journal of Education Policy. Vol. 15, No. 2, pp. 169-183.

Morris, E. (2001a) Professionalism and Trust - the future of teachers and teaching. Speech to the Social Market Foundation. November.

Morris, E. (2001b) Select Committee on Education and Skills. House of Commons. 24 October.

Morris, E. (2002) Transforming Education: The Middle Years. Speech to Demos. www.dfes.gov.uk. 21 March.

Morse, J and Lorsch, J . (1970) Beyond Theory Y. Harvard Business Review. May-June. pp. 61-68.

Mumford, A. (1995) The Learning Organisation in review. Industrial and Commercial Training. Vol. 27, No. 1.

Munro, L and Kimber, P. (1999) National Testing. In Bryce, T and Humes, W (Eds) Scottish Education, Edinburgh University Press. Edinburgh.

Munro, N. (1999) Schools get a rocket for ducking targets. Scottish Times Educational Supplement. 5 March.

Munro, R. (1995) Governing the new province of quality: autonomy, accounting and the dissemination of accountability. in A Wilkinson and H Willmott (eds). Making Quality Critical: New Perspectives on Organisational Change. Routledge. London.

Murgatroyd, S and Morgan, C. (1993) Total quality management and the school. Open University Press. Buckingham.

Murphy, R and Wilmut, J. (1999) Educational Assessment in an Era of Reform. In Demaine, J (Ed), Education Policy and Contemporary Politics, Macmillan. Basingstoke.

Myers, K. and Goldstein, H. (1997) Failing schools in a failing system. in Rethinking Educational change with heart and mind, Ed. A Hargreaves. ASCD Yearbook.

National Association of Head Teachers. (2000) National Numeracy Strategy Survey. NAHT. London.

National Audit Office. (2001) Inappropriate adjustments to NHS waiting lists. HC 452 Session 2001-2002. 19 December.

National College for School Leadership. (2002) Leadership development framework. www.ncsl.org.uk. .

Neill, S. (2002) National Curriculum Tests. National Union of Teachers. London, October.

New Shorter Oxford English Dictionary (1997) CD version

Nutley, S and Davies, H. (2000a) Making a Reality of Evidence-Based Practice: Some Lessons from the Diffusion of Innovation. Public Money & Management. Vol. 20, No. 4, pp. 35-42.

Nutley, S and Davies, H. (2000b) Developing learning organisations in the new NHS. British Medical Journal. Vol. 320: 998-1001.

Oakland, J. (1995) Total Quality Management. Butterworth-Heinemann. Oxford.

OECD. (2001) Knowledge and skills for life: first results from PISA 2000. OECD. Paris.

Office of Passenger Rail Franchising. (2000) On Track. OPRAF. www.opraf.gov.uk/ontrack/ontrack/1/overview.htm.

Office of the Deputy Prime Minister. (2002) Best Value Performance Indicators for 2001/2002. www.local-regions.odpm.gov.uk/bestvalue/indicators/pi2001-02/. London.

Office of the Information Commissioner. (1999) Case 98104 - The Sunday Times and others, and the Department of Education and Science. Dublin, 7 October. www.irlgov.ie.80/oic/2246_19a.htm.

Ofsted. (1998) School Evaluation Matters: The Business Excellence Model. www.ofsted.gov.uk/pubs/eval/annex/index.htm.

O'Neil, J. (1986) The disciplinary society: from Weber to Foucault. British Journal of Sociology. No. 37, pp. 42-60.

O'Neill, O. (2002) A Question of Trust. Reith Lectures. www.bbc.co.uk/radio4.

Organisation for Economic Co-operation and Development. (1995) Centre for Educational Research and Innovation: Schools Under Scrutiny. OECD. Paris.

Osborne, D. and Gaebler, T. (1992) Reinventing Government: How the Entrepreneurial Spirit is Transforming the Public Sector. Addison-Wesley. Reading.

Ostrom, V. (1974) The Intellectual Crisis in American Public Administration. University of Alabama Press. Alabama.

O'Sullivan. (1997) Learning Organisations - reengineering schools for life long learning. School Leadership and Management. Vol. 17, No. 2, pp. 217-230.

Oswald, A. (2001) An Economist's View of University League Tables. Public Money & Management. Vol. 21, No. 3, pp. 5-6.

Otley, D. (1978) Budget Use and Managerial Performance. Journal of Accounting Research. Spring, pp. 122-149.

Otley, D. (1980) The contingency theory of management accounting: achievement and prognosis. Accounting Organisations and Society. Vol. 5, No. 4, pp. 413-428.

Otley, D. (1999) Performance management: a framework for management control systems research. Management Accounting Research. Vol. 10, pp. 363-382.

Otley, D and Pollanen, R. (2000) Budgetary criteria in performance evaluation: a critical appraisal using new evidence. Accounting Organizations and Society. Vol. 25, pp. 483 - 496.

Ouchi, W. (1981) Theory Z: How American Business Can Meet the Japanese Challenge. Addison-Wesley. Reading M.A..

Ouston, J. (1999) School Effectiveness and School Improvement: Critique of a Movement. In Bush, T. etal (Eds) Educational Management: Redefining theory, policy and practice. Paul Chapman Publishing. London.

Ouston, J. (2000) Meeting the challenge of social disadvantage: managing school improvement. Paper - BEMAS seminar. Barnsley, June.

Ouston, J., Fidler, B and Earley, P. (1998) The Educational Accountability of Schools in England and Wales. Educational Policy. Vol. 12, No. 1 & 2, pp. 111-123.

Parsons, H. (1974) What Happened at Hawthorne?. Science. Vol. 183, 8 March.

Parsons, H. (1978) What caused the Hawthorne effect? A Scientific Detective Story. Administration and Society. Vol. 10, No. 3, pp. 259-282.

Passmore, B. (2002a) Happy teachers' only gripe is long hours. Times Educational Supplement. 1 February.

Passmore, B. (2002b) Labour standards drive fails to impress parents. Times Educational Supplement . 8 November.

Pedler, M., Burgoyne, J and Boydell, T. (1991) The Learning Company: a strategy for sustainable development. McGraw Hill. Maidenhead.

Performance and Innovation Unit. (2001) Better Policy Delivery and Design: A Discussion Paper. Cabinet Office. London.

Perrow, C. (1979) Complex organizations: a critical essay. Scott, Foresman and Co. London.

Peter, T and Waterman, R. (1982) In Search of Excellence. Macmillan. London.

Pfeffer, J. (1998) The Human Equation. Harvard Business School Press. Boston.

Philips, D. (2000) Beyond Travellers' Tales. Oxford Review of Education. Vol. 26, No. 1, pp. 49-62.

Plewis, I and Goldstein, H. (1998) The 1997 Education White Paper - a failure of standards. British Journal of Curriculum and Assessment. No. 8, pp. 17-20.

Plsek, P and Greenhalgh, T. (2001) The Challenge of complexity in health care. British Medical Journal. Vol. 323:625.

Plsek, P and Wilson, T. (2001) Complexity, leadership, and management in healthcare organisations. British Medical Journal. Vol. 323, pp. 746-749.

Pollard, A and Triggs, P. (2000) Policy, Practice and Pupil Experience. Continuum International Publishing Group. London.

Pollitt, C. (1993) Managerialism and the Public Services. Blackwell. Oxford.

Pollock, A., Shaoul, J and Vickers, N. (2002) Private finance and "value for money" in NHS hospitals: a policy in search of a rationale?. British Medical Journal. Vol. 324:1205-9.

Popper, K. (1963) Conjectures and Reflections. Routledge and Keagan Paul.

Powell, B. and Steelman, L. (1996) Bewitched, Bothered and Bewildering: The Use and Misuse of State SAT and ACT Scores. Harvard Educational Review. Vol. 66 No. 1 Spring pp. 27-55.

Power, S and Clark, A. (2000) The right to know: parents, school reports and parents' evenings. Research Papers in Education. Vol. 15, No. 1, pp. 25-48.

Power, S., Halpin, D., Whitty, G. (1997) Managing the state and the market: 'New' Education management in five countries. British Journal of Educational Studies. Vol. 45, No 4, pp 342-362.

Prescott, M. and Higgins, S. (1998) Labour MPs rebel at 'Napoleonic' regime. The Sunday Times. 12 July.

Pring, R. (2000) Editorial: Educational Research. British Journal of Educational Studies. Vol. 48, No. 1.

Pugh, D. (1990) Organization theory : selected readings . Penguin. London.

Pugh, D.S. and Hickson, D.J. (1989) Writers on Organisations. Penguin. London.

Pugh, D.S., Hickson, D.J and Hinings, C.R. (1971) Writers on Organisations. Penguin. London.

Qualifications and Curriculum Authority. (1999) Using Pupil Performance Information to Target School Improvement. QCA (99/432). London.

Quinlan, M and Scharaschkin, A (QCA). (1999) National curriculum testing: Problems and practicalities. Paper - British Educational Research Association, Conference. Brighton, September .

Quong, T and Walker, A. (1996) TQM and School Restructuring: a case study. School Organisation. Vol. 16, No. 2, pp. 219-231.

Raffe, D., Brannen, K., Croxford, L and Martin, C. (1999) Comparing England, Scotland, Wales and Northern Ireland: the Case for 'home internationals' in Comparative Research. Comparative Education . Vol. 35, No. 1, pp. 9-25.

Rafferty, F. (1999) Don't neglect the most able. Times Educational Supplement. 30 April.

Rafferty, F and Barnard, N. (1998) Blair warns teachers not to resist reforms. Times Educational Supplement. 2nd October.

Randall, J. (2001) Analysis: Where did Marconi go wrong? BBC . 5 July.

Rangone, A. (1997) Linking organizational effectiveness, key success factors and performance measures: an analytical framework. Management Accounting Research. Vol. 8, pp. 207-219.

Ranson, S. and Stewart, J. (1994) Management for the Public Domain: Enabling the Learning Society. Macmillan. London.

Regard. (2000) Redefining educational management. ESRC Award R45126449596. (www.regard.ac.uk - 10/12/00).

Rhodes, R. (1991) Introduction (To Hood 1991). Public Administration. Vol. 69, Spring, pp. 1-2.

Ribbins, P. (1999a) On Redefining Educational Management and Leadership. Educational Management and Administration. Vol. 27, No. 3, pp. 227 - 238.

Ribbins, P. (1999b) Setting up SCRELM. Educational Management and Administration. Vol. 27, No. 4, pp. 355-356.

Richards, C. (1988) Indicators and three types of educational monitoring systems: Implications for design.. Phi Delta Kappan. Vol. 69, No. 2. pp. 495-499.

Richardson, R. (1999) Performance Related Pay in Schools An assessment of the Green Papers. . National Union of Teachers. September.

Riley, K and Nuttall, D (Eds). (1994) Measuring Quality: Education Indicators - United Kingdom and International Perspectives. The Falmer Press. London.

Robson, W. (1994) Strategic Management and Information Systems. Pitman. London.

Roethlisberger, F and Dickson, W. (1964) Management and the Worker. John Wiley, New York.

Roslender, R. (1996) Relevance Lost and Found: Critical Perspectives on the Promise of Management Accounting. Critical Perspectives on Accounting. Vol. 7, pp. 533-561.

Sallis, E. (1993) Total Quality Management in education. Kogan Page. London.

Saunders, L. and Rudd, P. (1999) Schools' use of 'Value Added' Data: A science in the service of Art? Paper - British Educational Research Association Conference. Brighton, September.

Saunders, T. (2000) League Divisions. Junior Education. December.

Scottish Examination Board. (1993) 5-14 Assessment Unit: The Framework for National Testing. SEB. Dalkeith.

Senge, P. (1990) The Fifth Discipline: The Art and Practice of the Learning Organisation. Century Business. London.

Shadow Strategic Rail Authority. (2000) A new way of measuring performance. On Track. London.

Shafritz, J and Ott, J. (1992) Classics of Organization Theory. Books/Cole. Pacific Grove, CA.

Shaoul, J. (1999) Railpolitik: A Stakeholder Analysis of the Railways. Public Interest Report.. University of Manchester.

Shapiro, E. (1998) Fad Surfing in the Boardroom. Capstone. Oxford.

Shaw, L. (2002) Enron Unravelled. University of Durham Business School: Alimninews. Issue 3, June, Durham.

Shepard, L. (2000) The Role of Assessment in a Learning Culture. Educational Researcher. Vol. 29, No. 7, pp. 1-14.

Shepard, L. (2001) Interview with Lorrie Shepard - How to Fight A 'Death Star'. National Education Association Today Online. January, www.nea.org/neatoday.

Shields, J and Shields, M. (1998) Antecedents of Participative Budgeting. Accounting Organizations and Society. Vol. 23, No. 1, pp. 49-76.

Sizer, J. (1982) An insight into management accounting. Penguin. Harmondsworth.

Slater, J and Henry, J. (2001) 12-year olds to opt out of curriculum. Times Educational Supplement. 14 September.

Slavin, R. (1997) PET and the pendulum: Faddism in education and how to stop it. Phi Delta Kappan. No. 70, pp. 752-758.

Slavin, R. (2002) Evidence-Based Education Policies: Transforming Educational Practice and Research. Educational Researcher. Vol. 31, No. 7. pp. 15-21.

Smith, A. (2000) Public Service Productivity: Meeting the challenge. A Joint report. HM Treasury: Public Services Productivity Panel. .

Smith, L. (1998) Numeracy standards in schools slide further. The Times. 9 October.

Smith, P. (1993) Outcome -related performance indicators and organizational control in the public sector. British Journal of Management. Vol. 4. No. 3. pp. 135-152.

Smith, P. (1995a) On the unintended consequences of publishing performance data in the public sector. International Journal of Public Administration. Vol.18. No. 2 & 3. pp. 277-310.

Smith, P. (1995b) Performance indicators and control in the public sector. in Berry. A. Broadbent. J. Otley, D. (eds) Management Control. Macmillan. London.

Smithers, R. (2002) Schools cheat to boost exam results. The Guardian. 6 June.

Southworth, G. (2000) How Primary Schools Learn. Research Papers in Education. Vol 15. No. 3..

Spell, C. (1999) Where do management fashions come from, and how long do they stay?. Journal of Management History. Vol. 5. No. 6, pp. 334-384.

Spiegelhalter, D. (1999) Surgical audit: statistical lessons from Nightingale and Codman. Journal of Royal Statistical Society. Vol. 162. Part 1. pp. 45-58.

Spotiswood, C. (2000) Improving Police Performance: A New Approach to Measuring Police Efficiency. HM Treasury: Public Services Productivity Panel. 17 April.

Stacey, R. (1993) The chaos frontier: Creative strategic control for business. Butterworth-Heinemann. Oxford.

Starmer-Smith, C. (2002) Burglaries up as police focus on street crime. The Telegraph. 20 May.

Stata, R. (1989) Organizational Learning - The Key to Management Innovation. Sloan Management Review, . Spring 1989, P 63-74.

Stittle, J. (2002) Regulatory Control of the Track Access Charges of Railtrack plc. Public Money & Management. Vol. 22, No. 1, pp. 49-54.

Stobart, G. (2001) The Validity of National Curriculum Assessment. British Journal of Educational Studies. Vol. 49, No. 1, pp. 26-39.

Stout, J and Tymms, P. (1999) Science and the Literacy Hour. The Association for Science Education. Hatfield.

Sunday Times. (2001) The best value among the top 50 schools. The Sunday Times. 9 September.

Swieringa, R., Weick, K. (1987) Management Accounting and Action.. Accounting Organisations and Society. Vol 12, No 3, pp 293-308.

Talbot, C. (2000) Performing 'Performance' - A Comedy in Five Acts. Public Money & Management. Vol. 20, No. 4, pp. 63-68.

Taylor, F. (1916) The Principles of Scientific Management – Address to Cleveland Advertising Club . in Shafritz, J and Ott, J. (1992) Classics of Organization Theory. Wadsworth. Belmont.

Taylor, F. (1947) Scientific Management. Harper and Row. New York.

Teacher Training Agency. (1998) National Standards for Headteachers. TTA. London.

Telhaj, S. (2002) Dysfunctional Effects of Performance Indicators in English Education: A Principal-Agent Framework. Paper - British Educational Research Association, Conference. Exeter, September.

TES. (2002) Doubt over 'value-added' scores. Times Educational Supplement. 27 September.

TES Scotland. (2000) Joining up 10 and 14. Scottish Times Educational Supplement. 14 January.

The Economist. (1993) Return of the stopwatch. The Economist. 23 January.

The Telegraph. (1999) Scottish schools 'performing' better than English. The Telegraph. 26 November.

Theakston, K. (1998) New Labour, New Whitehall?. Public Policy and Administration. Vol. 13, No. 1. pp. 13-35.

Thietart, R (ed). (2001) Doing Management Research: a comprehensive guide. Sage. London.

Thompson, P. and McHugh, D. (2002) Work Organisations. Macmillan. London.

Thornton, K. (2001) League tables do not fit inclusion targets. Times Educational Supplement. 29 June.

Times. (1999) Surgeons' may drop high-risk patients'. The Times. 7 September.

Tomlinson, M. (2001) Giving more power to parents. The Sunday Times. 14 October.

Tooley, J. (2000) Reclaiming Education. Cassell. London.

Treasury. (2000) Spending Review 2000: New Public Spending Plans 2001-2004. HM Treasury. London.

Trivendi, B. (2002) Survey Reveals Geographic Illiteracy. http://news.nationalgeographic.com. .

Tymms, P. (1993) Accountability-can it be fair? Oxford Review of Education.. Vol. 19, No. 3. pp. 291-299.

Tymms, P. (1997) Errors in Multi-Level and OLS. In Fitz-Gibbon, C (Ed) The Value Added National Project: Final Report, School Curriculum and Assessment Authority. London.

Tymms, P. (1999) Baseline Assessment and Monitoring in Primary Schools: Achievements, Attitudes and Value-added Indicators. David Fulton. London.

Tymms, P and C. Fitz-Gibbon. (2001) Standards, Achievement and Educational Performance: A Cause for Celebration? In Phillips, R and Furlong, J (Eds), Education, Reform and the State: Twenty-Five Years of Politics, Policy and Practice., Routledge Falmer. London.

Vakola, M. (2000) Exploring the relationship between the use of evaluation in business process re-engineering and organisational learning and innovation. Journal of Management Development. Vol. 19, No. 10, pp. 812-835.

Van der Stede, W. (2000) The relationship between two consequences of budgetary controls: budgetary slack creation and managerial short-term orientation. Accounting Organizations and Society. Vol. 25, pp. 609-622.

Veasey, P. (1994) Managing a programme of business re-engineering in a diversified business. Long Range Planning. Vol. 27, No. 5, pp. 124-135.

Visscher, A., Karsten, S., de Jong, T. and Bosker, R. (1999) Evidence on the intended and unintended effects of publishing school performance indicators. Paper - E-BE-UK, 11-14 July. Durham.

Wallace, M. (2001) Really Good Ideas'?: Developing Evidence-Informed Policy and Practice in Educational Leadership and Management: A 'Rejoinder' to Rosalind Levacic and Ron Glatter. Educational Management and Administration. Vol. 29, No.1, pp. 27-34.

Wallace, M and McMahon. (1994) Planning for change in turbulent times : the case of multiracial primary schools. Cassell. London.

Waples. (2001) Big business learns the value of running struggling schools. The Sunday Times. 21 January.

Ward, H., Henry, J and Mansell, W. (2002) TES Primary Survey: Tests swallow up whole year. Times Educational Supplement . 1 November.

Waterhouse, R. (2001) Death rates reveal worst hospitals. The Sunday Times. 14 January.

Webster, D and Parsons, K. (1999) British Labour Party policy on educational selection 1996-8: a sociological analysis. Journal of Educational Policy. Vol. 14. No. 5. pp. 547-559.

Webster, P and Sherman, J. (1999) Blair 'scarred' by stubborn public sector. The Times. 7 July.

Weinstock, A. (1976) I Blame The Teachers. Times Educational Supplement. 23 January.

Welch, A. (1998) The Cult of Efficiency in Education: comparative reflections on the reality and the rhetoric. Comparative Education. Vol. 34. No. 2. pp. 157-175.

Weller, L. (1995) School restructuring and downsizing: using TQM to promote cost effectiveness. The TQM Magazine. Vol. 7. No. 6. pp. 11-16.

West, A and Pennell, H. (2000) Publishing School Examination Results in England: incentives and consequences. Educational Studies. Vol. 26, No. 4, pp. 423-436.

Wickens, P. (1995) The Ascendant Organisation. McMillan. London.

Wiggins, A and Tymms, P. (2002a) Dysfunctional effects of league tables: A comparison between English and Scottish primary schools. Public Money & Management. Vol. 22, No. 1, pp. 43-48.

Wiggins, A and Tymms, P. (2002b) Evidence-Based Policies: The Effect of Key Performance Indicators. Forthcoming.

Wiliam, D. (1994) Reconceptualising Validity, Dependability and Reliability for National Curriculum Assessment. In: Hutchison, D., and Schagen, I. (Eds.) How Reliable is National Curriculum Assessment? National Foundation for Educational Research. Slough.

Wiliam, D. (2001) Level Best? Levels of attainment in national curriculum assessment. Association of Teachers and Lecturers. London.

Wiliam, D. (2002) The meanings and consequences of educational assessments. Critical Quarterly. Vol. 42, No. 1, pp. 105-127.

Wilkinson, D. (1997) Whole system development - rethinking public service management. International Journal of Public Sector Management. Vol. 10, No. 7, pp. 505-533.

Willms, J. (1992) Monitoring School Performance: A guide for Educators. Falmer Press. London.

Wilson, R and Chua, W. (1993) Managerial Accounting: Method and Meaning. Chapman and Hall. London.

Witty, G., Power, S and Halpin, D. (1998) Devolution and Choice in Education. The School, the State and the Market. Open University Press. Buckingham.

Woodcock, B. (1998) Measuring Up: An Introduction to Theories and Concepts of Performance Measurement in Public Services. CIPFA. London.

Woods, P., Bagley, C and Glatter, R. (1998) School Choice and Competition: markets in the public interest. Routledge. London.

Woodward, T. (2002a) Edison's Failing Grade. Corpwatch, 20 June. www.corpwatch.org.

Woodward, W. (2002b) England still lags behind rest of UK in results. The Guardian. 22 August.